

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
6 December 2001 (06.12.2001)

PCT

(10) International Publication Number
WO 01/92581 A2

(51) International Patent Classification⁷: **C12Q 1/68**,
G01N 33/574, A61K 39/00, C12N 5/06, 5/08

(74) Agents: **POTTER, Jane, E., R.**; Seed Intellectual Prop-
erty Law Group PLLC, Suite 6300, 701 Fifth Avenue, Seat-
tle, WA 98104-7092 et al. (US).

(21) International Application Number: **PCT/US01/17756**

(22) International Filing Date: **29 May 2001 (29.05.2001)**

(25) Filing Language: **English**

(26) Publication Language: **English**

(30) Priority Data:
60/207,484 **26 May 2000 (26.05.2000)** **US**

(71) Applicant (for all designated States except US): **CORIXA
CORPORATION** [US/US]; Suite 200, 1124 Columbia
Street, Seattle, WA 98104 (US).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,
CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH,
GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC,
LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW,
MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK,
SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA,
ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM,
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,
IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF,
CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

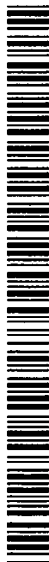
Published:

— without international search report and to be republished
upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.

(72) Inventors; and

(75) Inventors/Applicants (for US only): **ALGATE, Paul, A.**
[GB/US]; 580 Kalmia Place N.W., Issaquah, WA 98027
(US). **HARLOCKER, Susan, L.** [US/US]; 7522 13th
Avenue W., Seattle, WA 98117 (US). **JONES, Robert**
[GB/US]; 900 20th Avenue E., Seattle, WA 98112 (US).



WO 01/92581 A2

(54) Title: **COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS OF OVARIAN CANCER**

(57) Abstract: Compositions and methods for the therapy and diagnosis of cancer, particularly ovarian cancer, are disclosed. Illustrative compositions comprise one or more ovarian tumor polypeptides, immunogenic portions thereof, polynucleotides that encode such polypeptides, antigen presenting cell that expresses such polypeptides, and T cells that are specific for cells expressing such polypeptides. The disclosed compositions are useful, for example, in the diagnosis, prevention and/or treatment of diseases, particularly ovarian cancer.

COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS OF OVARIAN CANCER

STATEMENT REGARDING SEQUENCE LISTING

The Sequence Listing associated with this application is provided on
5 CD-ROM in lieu of a paper copy under AI § 801(a), and is hereby incorporated by
reference into the specification. Four CD-ROMs are provided containing identical
copies of the sequence listing: CD-ROM No. 1 is labeled "COPY 1 – SEQUENCE
LISTING PART," contains the file 497.app.txt which is 6.0 MB and created on May 29,
2001; CD-ROM No.2 is labeled "COPY 2 – SEQUENCE LISTING," contains the file
10 497.app.txt which is 6.0 MB and created on May 29, 2001; CD-ROM No. 3 is labeled
"COPY 3 – SEQUENCE LISTING PART," contains the file 497.app.txt which is 6.0
MB and created on May 29, 2001; CD-ROM No. 4 is labeled "CRF," contains the file
497.app.txt which is 6.0 Mb and created on May 29, 2001.

TECHNICAL FIELD OF THE INVENTION

15 The present invention relates generally to therapy and diagnosis of
cancer, such as ovarian cancer. The invention is more specifically related to
polypeptides, comprising at least a portion of an ovarian tumor protein, and to
polynucleotides encoding such polypeptides. Such polypeptides and polynucleotides
are useful in pharmaceutical compositions, *e.g.*, vaccines, and other compositions for
20 the diagnosis and treatment of ovarian cancer.

BACKGROUND OF THE INVENTION

Cancer is a significant health problem throughout the world. Although
advances have been made in detection and therapy of cancer, no vaccine or other
universally successful method for prevention and/or treatment is currently available.
25 Current therapies, which are generally based on a combination of chemotherapy or
surgery and radiation, continue to prove inadequate in many patients.

Ovarian cancer is a significant health problem for women in the United
States and throughout the world. Although advances have been made in detection and

therapy of this cancer, no vaccine or other universally successful method for prevention or treatment is currently available. Management of the disease currently relies on a combination of early diagnosis and aggressive treatment, which may include one or more of a variety of treatments such as surgery, radiotherapy, chemotherapy and
5 hormone therapy. The course of treatment for a particular cancer is often selected based on a variety of prognostic parameters, including an analysis of specific tumor markers. However, the use of established markers often leads to a result that is difficult to interpret, and high mortality continues to be observed in many cancer patients.

Immunotherapies have the potential to substantially improve cancer
10 treatment and survival. Such therapies may involve the generation or enhancement of an immune response to an ovarian carcinoma antigen. However, to date, relatively few ovarian carcinoma antigens are known and the generation of an immune response against such antigens has not been shown to be therapeutically beneficial.

Accordingly, there is a need in the art for improved methods for
15 identifying ovarian tumor antigens and for using such antigens in the therapy of ovarian cancer. The present invention fulfills these needs and further provides other related advantages.

In spite of considerable research into therapies for these and other
cancers, ovarian cancer remains difficult to diagnose and treat effectively. Accordingly,
20 there is a need in the art for improved methods for detecting and treating such cancers. The present invention fulfills these needs and further provides other related advantages.

SUMMARY OF THE INVENTION

In one aspect, the present invention provides polynucleotide compositions comprising a sequence selected from the group consisting of:

- 25 (a) sequences provided in SEQ ID NO: 1-10,912;
 (b) complements of the sequences provided in SEQ ID NO: 1-10,912;
 (c) sequences consisting of at least 20, 25, 30, 35, 40, 45, 50, 75 and 100 contiguous residues of a sequence provided in SEQ ID NO: 1-10,912;

(d) sequences that hybridize to a sequence provided in SEQ ID NO: 1-10,912, under moderate or highly stringent conditions;

(e) sequences having at least 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% identity to a sequence of SEQ ID NO: 1-10,912;

5 (f) degenerate variants of a sequence provided in SEQ ID NO: 1-10,912.

In one preferred embodiment, the polynucleotide compositions of the invention are expressed in at least about 20%, more preferably in at least about 30%,
10 and most preferably in at least about 50% of ovarian tumors samples tested, at a level that is at least about 2-fold, preferably at least about 5-fold, and most preferably at least about 10-fold higher than that for normal tissues.

The present invention, in another aspect, provides polypeptide compositions comprising an amino acid sequence that is encoded by a polynucleotide
15 sequence described above.

In certain preferred embodiments, the polypeptides and/or polynucleotides of the present invention are immunogenic, *i.e.*, they are capable of eliciting an immune response, particularly a humoral and/or cellular immune response, as further described herein.

20 The present invention further provides fragments, variants and/or derivatives of the disclosed polypeptide and/or polynucleotide sequences, wherein the fragments, variants and/or derivatives preferably have a level of immunogenic activity of at least about 50%, preferably at least about 70% and more preferably at least about 90% of the level of immunogenic activity of a polypeptide sequence encoded by a
25 polynucleotide sequence set forth in SEQ ID NO: 1-10,912.

The present invention further provides polynucleotides that encode a polypeptide described above, expression vectors comprising such polynucleotides and host cells transformed or transfected with such expression vectors.

30 Within other aspects, the present invention provides pharmaceutical compositions comprising a polypeptide or polynucleotide as described above and a physiologically acceptable carrier.

Within a related aspect of the present invention, the pharmaceutical compositions, *e.g.*, vaccine compositions, are provided for prophylactic or therapeutic applications. Such compositions generally comprise an immunogenic polypeptide or polynucleotide of the invention and an immunostimulant, such as an adjuvant.

5 The present invention further provides pharmaceutical compositions that comprise: (a) an antibody or antigen-binding fragment thereof that specifically binds to a polypeptide of the present invention, or a fragment thereof; and (b) a physiologically acceptable carrier.

 Within further aspects, the present invention provides pharmaceutical
10 compositions comprising: (a) an antigen presenting cell that expresses a polypeptide as described above and (b) a pharmaceutically acceptable carrier or excipient. Illustrative antigen presenting cells include dendritic cells, macrophages, monocytes, fibroblasts and B cells.

 Within related aspects, pharmaceutical compositions are provided that
15 comprise: (a) an antigen presenting cell that expresses a polypeptide as described above and (b) an immunostimulant.

 The present invention further provides, in other aspects, fusion proteins that comprise at least one polypeptide as described above, as well as polynucleotides encoding such fusion proteins, typically in the form of pharmaceutical compositions,
20 *e.g.*, vaccine compositions, comprising a physiologically acceptable carrier and/or an immunostimulant. The fusions proteins may comprise multiple immunogenic polypeptides or portions/variants thereof, as described herein, and may further comprise one or more polypeptide segments for facilitating the expression, purification and/or immunogenicity of the polypeptide(s).

25 Within further aspects, the present invention provides methods for stimulating an immune response in a patient, preferably a T cell response in a human patient, comprising administering a pharmaceutical composition described herein. The patient may be afflicted with ovarian cancer, in which case the methods provide treatment for the disease, or patient considered at risk for such a disease may be treated
30 prophylactically.

Within further aspects, the present invention provides methods for inhibiting the development of a cancer in a patient, comprising administering to a patient a pharmaceutical composition as recited above. The patient may be afflicted with ovarian cancer, in which case the methods provide treatment for the disease, or
5 patient considered at risk for such a disease may be treated prophylactically.

The present invention further provides, within other aspects, methods for removing tumor cells from a biological sample, comprising contacting a biological sample with T cells that specifically react with a polypeptide of the present invention, wherein the step of contacting is performed under conditions and for a time sufficient to
10 permit the removal of cells expressing the protein from the sample.

Within related aspects, methods are provided for inhibiting the development of a cancer in a patient, comprising administering to a patient a biological sample treated as described above.

Methods are further provided, within other aspects, for stimulating
15 and/or expanding T cells specific for a polypeptide of the present invention, comprising contacting T cells with one or more of: (i) a polypeptide as described above; (ii) a polynucleotide encoding such a polypeptide; and/or (iii) an antigen presenting cell that expresses such a polypeptide; under conditions and for a time sufficient to permit the stimulation and/or expansion of T cells. Isolated T cell populations comprising T cells
20 prepared as described above are also provided.

Within further aspects, the present invention provides methods for inhibiting the development of a cancer in a patient, comprising administering to a patient an effective amount of a T cell population as described above.

The present invention further provides methods for inhibiting the
25 development of a cancer in a patient, comprising the steps of: (a) incubating CD4⁺ and/or CD8⁺ T cells isolated from a patient with one or more of: (i) a polypeptide comprising at least an immunogenic portion of polypeptide disclosed herein; (ii) a polynucleotide encoding such a polypeptide; and (iii) an antigen-presenting cell that expressed such a polypeptide; and (b) administering to the patient an effective amount
30 of the proliferated T cells, and thereby inhibiting the development of a cancer in the

patient. Proliferated cells may, but need not, be cloned prior to administration to the patient.

Within further aspects, the present invention provides methods for determining the presence or absence of a cancer, preferably an ovarian cancer, in a patient comprising: (a) contacting a biological sample obtained from a patient with a binding agent that binds to a polypeptide as recited above; (b) detecting in the sample an amount of polypeptide that binds to the binding agent; and (c) comparing the amount of polypeptide with a predetermined cut-off value, and therefrom determining the presence or absence of a cancer in the patient. Within preferred embodiments, the binding agent is an antibody, more preferably a monoclonal antibody.

The present invention also provides, within other aspects, methods for monitoring the progression of a cancer in a patient. Such methods comprise the steps of: (a) contacting a biological sample obtained from a patient at a first point in time with a binding agent that binds to a polypeptide as recited above; (b) detecting in the sample an amount of polypeptide that binds to the binding agent; (c) repeating steps (a) and (b) using a biological sample obtained from the patient at a subsequent point in time; and (d) comparing the amount of polypeptide detected in step (c) with the amount detected in step (b) and therefrom monitoring the progression of the cancer in the patient.

The present invention further provides, within other aspects, methods for determining the presence or absence of a cancer in a patient, comprising the steps of: (a) contacting a biological sample, e.g., tumor sample, serum sample, etc., obtained from a patient with an oligonucleotide that hybridizes to a polynucleotide that encodes a polypeptide of the present invention; (b) detecting in the sample a level of a polynucleotide, preferably mRNA, that hybridizes to the oligonucleotide; and (c) comparing the level of polynucleotide that hybridizes to the oligonucleotide with a predetermined cut-off value, and therefrom determining the presence or absence of a cancer in the patient. Within certain embodiments, the amount of mRNA is detected via polymerase chain reaction using, for example, at least one oligonucleotide primer that hybridizes to a polynucleotide encoding a polypeptide as recited above, or a complement of such a polynucleotide. Within other embodiments, the amount of

mRNA is detected using a hybridization technique, employing an oligonucleotide probe that hybridizes to a polynucleotide that encodes a polypeptide as recited above, or a complement of such a polynucleotide.

In related aspects, methods are provided for monitoring the progression
5 of a cancer in a patient, comprising the steps of: (a) contacting a biological sample obtained from a patient with an oligonucleotide that hybridizes to a polynucleotide that encodes a polypeptide of the present invention; (b) detecting in the sample an amount of a polynucleotide that hybridizes to the oligonucleotide; (c) repeating steps (a) and (b) using a biological sample obtained from the patient at a subsequent point in time; and
10 (d) comparing the amount of polynucleotide detected in step (c) with the amount detected in step (b) and therefrom monitoring the progression of the cancer in the patient.

Within further aspects, the present invention provides antibodies, such as monoclonal antibodies, that bind to a polypeptide as described above, as well as
15 diagnostic kits comprising such antibodies. Diagnostic kits comprising one or more oligonucleotide probes or primers as described above are also provided.

These and other aspects of the present invention will become apparent upon reference to the following detailed description. All references disclosed herein are hereby incorporated by reference in their entirety as if each was incorporated
20 individually.

BRIEF DESCRIPTION OF THE SEQUENCE IDENTIFIERS

SEQ ID NO: 1 represents the cDNA sequence for clone T29202.
SEQ ID NO: 2 represents the cDNA sequence for clone T29204.
SEQ ID NO: 3 represents the cDNA sequence for clone T29205.
25 SEQ ID NO: 4 represents the cDNA sequence for clone T29208.
SEQ ID NO: 5 represents the cDNA sequence for clone T29210.
SEQ ID NO: 6 represents the cDNA sequence for clone T29221.
SEQ ID NO: 7 represents the cDNA sequence for clone T29223.
SEQ ID NO: 8 represents the cDNA sequence for clone T34666.
30 SEQ ID NO: 9 represents the cDNA sequence for clone T34668.

SEQ ID NO: 4102 represents the cDNA sequence for clone AA410701.
SEQ ID NO: 4103 represents the cDNA sequence for clone AA410703.
SEQ ID NO: 4104 represents the cDNA sequence for clone AA410687.
SEQ ID NO: 4105 represents the cDNA sequence for clone AA410719.
5 SEQ ID NO: 4106 represents the cDNA sequence for clone AA410720.
SEQ ID NO: 4107 represents the cDNA sequence for clone AA410721.
SEQ ID NO: 4108 represents the cDNA sequence for clone AA410722.
SEQ ID NO: 4109 represents the cDNA sequence for clone AA410732.
SEQ ID NO: 4110 represents the cDNA sequence for clone AA410734.
10 SEQ ID NO: 4111 represents the cDNA sequence for clone AA410740.
SEQ ID NO: 4112 represents the cDNA sequence for clone AA410627.
SEQ ID NO: 4113 represents the cDNA sequence for clone AA410542.
SEQ ID NO: 4114 represents the cDNA sequence for clone AA410549.
SEQ ID NO: 4115 represents the cDNA sequence for clone AA410749.
15 SEQ ID NO: 4116 represents the cDNA sequence for clone AA410769.
SEQ ID NO: 4117 represents the cDNA sequence for clone AA410776.
SEQ ID NO: 4118 represents the cDNA sequence for clone AA410778.
SEQ ID NO: 4119 represents the cDNA sequence for clone AA410786.
SEQ ID NO: 4120 represents the cDNA sequence for clone AA410792.
20 SEQ ID NO: 4121 represents the cDNA sequence for clone AA410804.
SEQ ID NO: 4122 represents the cDNA sequence for clone AA410699.
SEQ ID NO: 4123 represents the cDNA sequence for clone AA410761.
SEQ ID NO: 4124 represents the cDNA sequence for clone AA410832.
SEQ ID NO: 4125 represents the cDNA sequence for clone AA410834.
25 SEQ ID NO: 4126 represents the cDNA sequence for clone AA410835.
SEQ ID NO: 4127 represents the cDNA sequence for clone AA410838.
SEQ ID NO: 4128 represents the cDNA sequence for clone AA410849.
SEQ ID NO: 4129 represents the cDNA sequence for clone AA410853.
SEQ ID NO: 4130 represents the cDNA sequence for clone AA410860.
30 SEQ ID NO: 4131 represents the cDNA sequence for clone AA410864.
SEQ ID NO: 4132 represents the cDNA sequence for clone AA410820.

SEQ ID NO: 6117 represents the cDNA sequence for clone AA481375.
SEQ ID NO: 6118 represents the cDNA sequence for clone AA481384.
SEQ ID NO: 6119 represents the cDNA sequence for clone AA481448.
SEQ ID NO: 6120 represents the cDNA sequence for clone AA481451.
5 SEQ ID NO: 6121 represents the cDNA sequence for clone AA481453.
SEQ ID NO: 6122 represents the cDNA sequence for clone AA481454.
SEQ ID NO: 6123 represents the cDNA sequence for clone AA481455.
SEQ ID NO: 6124 represents the cDNA sequence for clone AA481456.
SEQ ID NO: 6125 represents the cDNA sequence for clone AA481467.
10 SEQ ID NO: 6126 represents the cDNA sequence for clone AA481476.
SEQ ID NO: 6127 represents the cDNA sequence for clone AA481482.
SEQ ID NO: 6128 represents the cDNA sequence for clone AA481483.
SEQ ID NO: 6129 represents the cDNA sequence for clone AA481687.
SEQ ID NO: 6130 represents the cDNA sequence for clone AA481689.
15 SEQ ID NO: 6131 represents the cDNA sequence for clone AA481690.
SEQ ID NO: 6132 represents the cDNA sequence for clone AA481691.
SEQ ID NO: 6133 represents the cDNA sequence for clone AA481692.
SEQ ID NO: 6134 represents the cDNA sequence for clone AA481693.
SEQ ID NO: 6135 represents the cDNA sequence for clone AA481703.
20 SEQ ID NO: 6136 represents the cDNA sequence for clone AA481709.
SEQ ID NO: 6137 represents the cDNA sequence for clone AA481714.
SEQ ID NO: 6138 represents the cDNA sequence for clone AA481717.
SEQ ID NO: 6139 represents the cDNA sequence for clone AA481719.
SEQ ID NO: 6140 represents the cDNA sequence for clone AA481720.
25 SEQ ID NO: 6141 represents the cDNA sequence for clone AA481721.
SEQ ID NO: 6142 represents the cDNA sequence for clone AA481744.
SEQ ID NO: 6143 represents the cDNA sequence for clone AA481753.
SEQ ID NO: 6144 represents the cDNA sequence for clone AA481754.
SEQ ID NO: 6145 represents the cDNA sequence for clone AA481757.
30 SEQ ID NO: 6146 represents the cDNA sequence for clone AA485351.
SEQ ID NO: 6147 represents the cDNA sequence for clone AA485360.

- SEQ ID NO: 10902 represents the DNA sequence for clone O1074C.
SEQ ID NO: 10903 represents an extended sequence for clone O1074C.
SEQ ID NO: 10904 represents the DNA sequence for clone O1075C.
SEQ ID NO: 10905 represents an extended sequence for clone O1075C.
5 SEQ ID NO: 10906 represents the DNA sequence for clone O1076C.
SEQ ID NO: 10907 represents an extended sequence for clone O1076C.
SEQ ID NO: 10908 represents the DNA sequence for clone O1077C.
SEQ ID NO: 10909 represents the DNA sequence for clone O1078C.
SEQ ID NO: 10910 represents an extended sequence for clone O1078C.
10 SEQ ID NO: 10911 represents the DNA sequence for clone O1079C.
SEQ ID NO: 10912 represents an extended sequence for clone O1079C.

DETAILED DESCRIPTION OF THE INVENTION

The present invention is directed generally to compositions and their use in the therapy and diagnosis of cancer, particularly ovarian cancer. As described further
15 below, illustrative compositions of the present invention include, but are not restricted to, polypeptides, particularly immunogenic polypeptides, polynucleotides encoding such polypeptides, antibodies and other binding agents, antigen presenting cells (APCs) and immune system cells (*e.g.*, T cells).

The practice of the present invention will employ, unless indicated
20 specifically to the contrary, conventional methods of virology, immunology, microbiology, molecular biology and recombinant DNA techniques within the skill of the art, many of which are described below for the purpose of illustration. Such techniques are explained fully in the literature. See, *e.g.*, Sambrook, et al. Molecular Cloning: A Laboratory Manual (2nd Edition, 1989); Maniatis et al. Molecular Cloning:
25 A Laboratory Manual (1982); DNA Cloning: A Practical Approach, vol. I & II (D. Glover, ed.); Oligonucleotide Synthesis (N. Gait, ed., 1984); Nucleic Acid Hybridization (B. Hames & S. Higgins, eds., 1985); Transcription and Translation (B. Hames & S. Higgins, eds., 1984); Animal Cell Culture (R. Freshney, ed., 1986); Perbal, A Practical Guide to Molecular Cloning (1984).

All publications, patents and patent applications cited herein, whether supra or infra, are hereby incorporated by reference in their entirety.

As used in this specification and the appended claims, the singular forms "a," "an" and "the" include plural references unless the content clearly dictates
5 otherwise.

Polypeptide Compositions

As used herein, the term "polypeptide" is used in its conventional meaning, *i.e.*, as a sequence of amino acids. The polypeptides are not limited to a specific length of the product; thus, peptides, oligopeptides, and proteins are included
10 within the definition of polypeptide, and such terms may be used interchangeably herein unless specifically indicated otherwise. This term also does not refer to or exclude post-expression modifications of the polypeptide, for example, glycosylations, acetylations, phosphorylations and the like, as well as other modifications known in the art, both
15 naturally occurring and non-naturally occurring. A polypeptide may be an entire protein, or a subsequence thereof. Particular polypeptides of interest in the context of this invention are amino acid subsequences comprising epitopes, *i.e.*, antigenic determinants substantially responsible for the immunogenic properties of a polypeptide and being capable of evoking an immune response.

Particularly illustrative polypeptides of the present invention comprise
20 those encoded by a polynucleotide sequence set forth in any one of SEQ ID NO: 1-10,912, or a sequence that hybridizes under moderately stringent conditions, or, alternatively, under highly stringent conditions, to a polynucleotide sequence set forth in any one of SEQ ID NO: 1-10,912.

The polypeptides of the present invention are sometimes herein referred
25 to as ovarian tumor proteins or ovarian tumor polypeptides, as an indication that their identification has been based at least in part upon their increased levels of expression in ovarian tumor samples. Thus, an "ovarian tumor polypeptide" or "ovarian tumor protein," refers generally to a polypeptide sequence of the present invention, or a polynucleotide sequence encoding such a polypeptide, that is expressed in a substantial
30 proportion of ovarian tumor samples, for example preferably greater than about 20%.

more preferably greater than about 30%, and most preferably greater than about 50% or more of ovarian tumor samples tested, at a level that is at least two fold, and preferably at least five fold, greater than the level of expression in normal tissues, as determined using a representative assay provided herein. An ovarian tumor polypeptide sequence
5 of the invention, based upon its increased level of expression in tumor cells, has particular utility both as a diagnostic marker as well as a therapeutic target, as further described below.

In certain preferred embodiments, the polypeptides of the invention are immunogenic, *i.e.*, they react detectably within an immunoassay (such as an ELISA or
10 T-cell stimulation assay) with antisera and/or T-cells from a patient with ovarian cancer. Screening for immunogenic activity can be performed using techniques well known to the skilled artisan. For example, such screens can be performed using methods such as those described in Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988. In one illustrative example, a polypeptide may be
15 immobilized on a solid support and contacted with patient sera to allow binding of antibodies within the sera to the immobilized polypeptide. Unbound sera may then be removed and bound antibodies detected using, for example, ¹²⁵I-labeled Protein A.

As would be recognized by the skilled artisan, immunogenic portions of the polypeptides disclosed herein are also encompassed by the present invention. An
20 "immunogenic portion," as used herein, is a fragment of an immunogenic polypeptide of the invention that itself is immunologically reactive (*i.e.*, specifically binds) with the B-cells and/or T-cell surface antigen receptors that recognize the polypeptide. Immunogenic portions may generally be identified using well known techniques, such as those summarized in Paul, *Fundamental Immunology*, 3rd ed., 243-247 (Raven Press,
25 1993) and references cited therein. Such techniques include screening polypeptides for the ability to react with antigen-specific antibodies, antisera and/or T-cell lines or clones. As used herein, antisera and antibodies are "antigen-specific" if they specifically bind to an antigen (*i.e.*, they react with the protein in an ELISA or other immunoassay, and do not react detectably with unrelated proteins). Such antisera and
30 antibodies may be prepared as described herein, and using well-known techniques.

In one preferred embodiment, an immunogenic portion of a polypeptide of the present invention is a portion that reacts with antisera and/or T-cells at a level that is not substantially less than the reactivity of the full-length polypeptide (*e.g.*, in an ELISA and/or T-cell reactivity assay). Preferably, the level of immunogenic activity of the immunogenic portion is at least about 50%, preferably at least about 70% and most preferably greater than about 90% of the immunogenicity for the full-length polypeptide. In some instances, preferred immunogenic portions will be identified that have a level of immunogenic activity greater than that of the corresponding full-length polypeptide, *e.g.*, having greater than about 100% or 150% or more immunogenic activity.

In certain other embodiments, illustrative immunogenic portions may include peptides in which an N-terminal leader sequence and/or transmembrane domain have been deleted. Other illustrative immunogenic portions will contain a small N- and/or C-terminal deletion (*e.g.*, 1-30 amino acids, preferably 5-15 amino acids), relative to the mature protein.

In another embodiment, a polypeptide composition of the invention may also comprise one or more polypeptides that are immunologically reactive with T cells and/or antibodies generated against a polypeptide of the invention, particularly a polypeptide having an amino acid sequence disclosed herein, or to an immunogenic fragment or variant thereof.

In another embodiment of the invention, polypeptides are provided that comprise one or more polypeptides that are capable of eliciting T cells and/or antibodies that are immunologically reactive with one or more polypeptides described herein, or one or more polypeptides encoded by contiguous nucleic acid sequences contained in the polynucleotide sequences disclosed herein, or immunogenic fragments or variants thereof, or to one or more nucleic acid sequences which hybridize to one or more of these sequences under conditions of moderate to high stringency.

The present invention, in another aspect, provides polypeptide fragments comprising at least about 5, 10, 15, 20, 25, 50, or 100 contiguous amino acids, or more, including all intermediate lengths, of a polypeptide compositions set forth herein, such

as those encoded by a polynucleotide sequence set forth in a sequence of SEQ ID NO: 1-10,912.

In another aspect, the present invention provides variants of the polypeptide compositions described herein. Polypeptide variants generally encompassed by the present invention will typically exhibit at least about 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% or more identity (determined as described below), along its length, to a polypeptide sequences set forth herein.

In one preferred embodiment, the polypeptide fragments and variants provided by the present invention are immunologically reactive with an antibody and/or T-cell that reacts with a full-length polypeptide specifically set forth herein.

In another preferred embodiment, the polypeptide fragments and variants provided by the present invention exhibit a level of immunogenic activity of at least about 50%, preferably at least about 70%, and most preferably at least about 90% or more of that exhibited by a full-length polypeptide sequence specifically set forth herein.

A polypeptide "variant," as the term is used herein, is a polypeptide that typically differs from a polypeptide specifically disclosed herein in one or more substitutions, deletions, additions and/or insertions. Such variants may be naturally occurring or may be synthetically generated, for example, by modifying one or more of the above polypeptide sequences of the invention and evaluating their immunogenic activity as described herein and/or using any of a number of techniques well known in the art.

For example, certain illustrative variants of the polypeptides of the invention include those in which one or more portions, such as an N-terminal leader sequence or transmembrane domain, have been removed. Other illustrative variants include variants in which a small portion (*e.g.*, 1-30 amino acids, preferably 5-15 amino acids) has been removed from the N- and/or C-terminal of the mature protein.

In many instances, a variant will contain conservative substitutions. A "conservative substitution" is one in which an amino acid is substituted for another amino acid that has similar properties, such that one skilled in the art of peptide

chemistry would expect the secondary structure and hydrophobic nature of the polypeptide to be substantially unchanged. As described above, modifications may be made in the structure of the polynucleotides and polypeptides of the present invention and still obtain a functional molecule that encodes a variant or derivative polypeptide with desirable characteristics, *e.g.*, with immunogenic characteristics. When it is desired to alter the amino acid sequence of a polypeptide to create an equivalent, or even an improved, immunogenic variant or portion of a polypeptide of the invention, one skilled in the art will typically change one or more of the codons of the encoding DNA sequence according to Table 1.

For example, certain amino acids may be substituted for other amino acids in a protein structure without appreciable loss of interactive binding capacity with structures such as, for example, antigen-binding regions of antibodies or binding sites on substrate molecules. Since it is the interactive capacity and nature of a protein that defines that protein's biological functional activity, certain amino acid sequence substitutions can be made in a protein sequence, and, of course, its underlying DNA coding sequence, and nevertheless obtain a protein with like properties. It is thus contemplated that various changes may be made in the peptide sequences of the disclosed compositions, or corresponding DNA sequences which encode said peptides without appreciable loss of their biological utility or activity.

20

TABLE 1

Amino Acids			Codons						
Alanine	Ala	A	GCA	GCC	GCG	GCU			
Cysteine	Cys	C	UGC	UGU					
Aspartic acid	Asp	D	GAC	GAU					
Glutamic acid	Glu	E	GAA	GAG					
Phenylalanine	Phe	F	UUC	UUU					
Glycine	Gly	G	GGA	GGC	GGG	GGU			
Histidine	His	H	CAC	CAU					
Isoleucine	Ile	I	AUA	AUC	AUU				
Lysine	Lys	K	AAA	AAG					
Leucine	Leu	L	UUA	UUG	CUA	CUC	CUG	CUU	
Methionine	Met	M	AUG						
Asparagine	Asn	N	AAC	AAU					
Proline	Pro	P	CCA	CCC	CCG	CCU			
Glutamine	Gln	Q	CAA	CAG					
Arginine	Arg	R	AGA	AGG	CGA	CGC	CGG	CGU	
Serine	Ser	S	AGC	AGU	UCA	UCC	UCG	UCU	
Threonine	Thr	T	ACA	ACC	ACG	ACU			
Valine	Val	V	GUA	GUC	GUG	GUU			
Tryptophan	Trp	W	UGG						
Tyrosine	Tyr	Y	UAC	UAU					

In making such changes, the hydropathic index of amino acids may be considered. The importance of the hydropathic amino acid index in conferring interactive biologic function on a protein is generally understood in the art (Kyte and Doolittle, 1982, incorporated herein by reference). It is accepted that the relative hydropathic character of the amino acid contributes to the secondary structure of the resultant protein, which in turn defines the interaction of the protein with other molecules, for example, enzymes, substrates, receptors, DNA, antibodies, antigens, and the like. Each amino acid has been assigned a hydropathic index on the basis of its hydrophobicity and charge characteristics (Kyte and Doolittle, 1982). These values are:

isoleucine (+4.5); valine (+4.2); leucine (+3.8); phenylalanine (+2.8); cysteine/cystine (+2.5); methionine (+1.9); alanine (+1.8); glycine (-0.4); threonine (-0.7); serine (-0.8); tryptophan (-0.9); tyrosine (-1.3); proline (-1.6); histidine (-3.2); glutamate (-3.5); glutamine (-3.5); aspartate (-3.5); asparagine (-3.5); lysine (-3.9); and arginine (-4.5).

5 It is known in the art that certain amino acids may be substituted by other amino acids having a similar hydropathic index or score and still result in a protein with similar biological activity, *i.e.* still obtain a biological functionally equivalent protein. In making such changes, the substitution of amino acids whose hydropathic indices are within ± 2 is preferred, those within ± 1 are particularly preferred, and those within ± 0.5
10 are even more particularly preferred. It is also understood in the art that the substitution of like amino acids can be made effectively on the basis of hydrophilicity. U. S. Patent 4,554,101 (specifically incorporated herein by reference in its entirety), states that the greatest local average hydrophilicity of a protein, as governed by the hydrophilicity of its adjacent amino acids, correlates with a biological property of the protein.

15 As detailed in U. S. Patent 4,554,101, the following hydrophilicity values have been assigned to amino acid residues: arginine (+3.0); lysine (+3.0); aspartate (+3.0 \pm 1); glutamate (+3.0 \pm 1); serine (+0.3); asparagine (+0.2); glutamine (+0.2); glycine (0); threonine (-0.4); proline (-0.5 \pm 1); alanine (-0.5); histidine (-0.5); cysteine (-1.0); methionine (-1.3); valine (-1.5); leucine (-1.8); isoleucine (-1.8); tyrosine (-
20 2.3); phenylalanine (-2.5); tryptophan (-3.4). It is understood that an amino acid can be substituted for another having a similar hydrophilicity value and still obtain a biologically equivalent, and in particular, an immunologically equivalent protein. In such changes, the substitution of amino acids whose hydrophilicity values are within ± 2 is preferred, those within ± 1 are particularly preferred, and those within ± 0.5 are even
25 more particularly preferred.

 As outlined above, amino acid substitutions are generally therefore based on the relative similarity of the amino acid side-chain substituents, for example, their hydrophobicity, hydrophilicity, charge, size, and the like. Exemplary substitutions that take various of the foregoing characteristics into consideration are well known to those
30 of skill in the art and include: arginine and lysine; glutamate and aspartate; serine and threonine; glutamine and asparagine; and valine, leucine and isoleucine.

In addition, any polynucleotide may be further modified to increase stability *in vivo*. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends; the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages in the backbone; and/or the inclusion of
5 nontraditional bases such as inosine, queosine and wybutosine, as well as acetyl-methyl-, thio- and other modified forms of adenine, cytidine, guanine, thymine and uridine.

Amino acid substitutions may further be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity and/or the amphipathic
10 nature of the residues. For example, negatively charged amino acids include aspartic acid and glutamic acid; positively charged amino acids include lysine and arginine; and amino acids with uncharged polar head groups having similar hydrophilicity values include leucine, isoleucine and valine; glycine and alanine; asparagine and glutamine; and serine, threonine, phenylalanine and tyrosine. Other groups of amino acids that may
15 represent conservative changes include: (1) ala, pro, gly, glu, asp, gln, asn, ser, thr; (2) cys, ser, tyr, thr; (3) val, ile, leu, met, ala, phe; (4) lys, arg, his; and (5) phe, tyr, trp, his. A variant may also, or alternatively, contain nonconservative changes. In a preferred embodiment, variant polypeptides differ from a native sequence by substitution, deletion or addition of five amino acids or fewer. Variants may also (or
20 alternatively) be modified by, for example, the deletion or addition of amino acids that have minimal influence on the immunogenicity, secondary structure and hydrophobic nature of the polypeptide.

As noted above, polypeptides may comprise a signal (or leader) sequence at the N-terminal end of the protein, which co-translationally or post-translationally
25 directs transfer of the protein. The polypeptide may also be conjugated to a linker or other sequence for ease of synthesis, purification or identification of the polypeptide (e.g., poly-His), or to enhance binding of the polypeptide to a solid support. For example, a polypeptide may be conjugated to an immunoglobulin Fc region.

When comparing polypeptide sequences, two sequences are said to be
30 "identical" if the sequence of amino acids in the two sequences is the same when aligned for maximum correspondence, as described below. Comparisons between two

sequences are typically performed by comparing the sequences over a comparison window to identify and compare local regions of sequence similarity. A "comparison window" as used herein, refers to a segment of at least about 20 contiguous positions, usually 30 to about 75, 40 to about 50, in which a sequence may be compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned.

Optimal alignment of sequences for comparison may be conducted using the Megalign program in the Lasergene suite of bioinformatics software (DNASTAR, Inc., Madison, WI), using default parameters. This program embodies several alignment schemes described in the following references: Dayhoff, M.O. (1978) A model of evolutionary change in proteins – Matrices for detecting distant relationships. In Dayhoff, M.O. (ed.) *Atlas of Protein Sequence and Structure*, National Biomedical Research Foundation, Washington DC Vol. 5, Suppl. 3, pp. 345-358; Hein J. (1990) Unified Approach to Alignment and Phylogenies pp. 626-645 *Methods in Enzymology* vol. 183, Academic Press, Inc., San Diego, CA; Higgins, D.G. and Sharp, P.M. (1989) *CABIOS* 5:151-153; Myers, E.W. and Muller W. (1988) *CABIOS* 4:11-17; Robinson, E.D. (1971) *Comb. Theor* 11:105; Saitou, N. Nei, M. (1987) *Mol. Biol. Evol.* 4:406-425; Sneath, P.H.A. and Sokal, R.R. (1973) *Numerical Taxonomy – the Principles and Practice of Numerical Taxonomy*, Freeman Press, San Francisco, CA; Wilbur, W.J. and Lipman, D.J. (1983) *Proc. Natl. Acad., Sci. USA* 80:726-730.

Alternatively, optimal alignment of sequences for comparison may be conducted by the local identity algorithm of Smith and Waterman (1981) *Add. APL. Math* 2:482, by the identity alignment algorithm of Needleman and Wunsch (1970) *J. Mol. Biol.* 48:443, by the search for similarity methods of Pearson and Lipman (1988) *Proc. Natl. Acad. Sci. USA* 85: 2444, by computerized implementations of these algorithms (GAP, BESTFIT, BLAST, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group (GCG), 575 Science Dr., Madison, WI), or by inspection.

One preferred example of algorithms that are suitable for determining percent sequence identity and sequence similarity are the BLAST and BLAST 2.0 algorithms, which are described in Altschul et al. (1977) *Nucl. Acids Res.* 25:3389-3402

and Altschul et al. (1990) *J. Mol. Biol.* 215:403-410, respectively. BLAST and BLAST 2.0 can be used, for example with the parameters described herein, to determine percent sequence identity for the polynucleotides and polypeptides of the invention. Software for performing BLAST analyses is publicly available through the National Center for
5 Biotechnology Information. For amino acid sequences, a scoring matrix can be used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is
10 reached. The BLAST algorithm parameters W, T and X determine the sensitivity and speed of the alignment.

In one preferred approach, the "percentage of sequence identity" is determined by comparing two optimally aligned sequences over a window of comparison of at least 20 positions, wherein the portion of the polypeptide sequence in
15 the comparison window may comprise additions or deletions (*i.e.*, gaps) of 20 percent or less, usually 5 to 15 percent, or 10 to 12 percent, as compared to the reference sequences (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical amino acid residue occurs in both sequences to yield the number of
20 matched positions, dividing the number of matched positions by the total number of positions in the reference sequence (*i.e.*, the window size) and multiplying the results by 100 to yield the percentage of sequence identity.

Within other illustrative embodiments, a polypeptide may be a xenogeneic polypeptide that comprises an polypeptide having substantial sequence
25 identity, as described above, to the human polypeptide (also termed autologous antigen) which served as a reference polypeptide, but which xenogeneic polypeptide is derived from a different, non-human species. One skilled in the art will recognize that "self" antigens are often poor stimulators of CD8+ and CD4+ T-lymphocyte responses, and therefore efficient immunotherapeutic strategies directed against tumor
30 polypeptides require the development of methods to overcome immune tolerance to particular self tumor polypeptides. For example, humans immunized with prostate

protein from a xenogeneic (non human) origin are capable of mounting an immune response against the counterpart human protein, *e.g.* the human prostate tumor protein present on human tumor cells. Accordingly, the present invention provides methods for purifying the xenogeneic form of the tumor proteins set forth herein, such as those
5 encoded by polynucleotide sequences set forth in SEQ ID NO: 1-10,912.

Therefore, one aspect of the present invention provides xenogeneic variants of the polypeptide compositions described herein. Such xenogeneic variants generally encompassed by the present invention will typically exhibit at least about 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% or
10 more identity along their lengths, to a polypeptide sequences set forth herein.

More particularly, the invention is directed to mouse, rat, monkey, porcine and other non-human polypeptides which can be used as xenogeneic forms of human polypeptides set forth herein, to induce immune responses directed against tumor polypeptides of the invention.

15 Within other illustrative embodiments, a polypeptide may be a fusion polypeptide that comprises multiple polypeptides as described herein, or that comprises at least one polypeptide as described herein and an unrelated sequence, such as a known tumor protein. A fusion partner may, for example, assist in providing T helper epitopes (an immunological fusion partner), preferably T helper epitopes recognized by humans,
20 or may assist in expressing the protein (an expression enhancer) at higher yields than the native recombinant protein. Certain preferred fusion partners are both immunological and expression enhancing fusion partners. Other fusion partners may be selected so as to increase the solubility of the polypeptide or to enable the polypeptide to be targeted to desired intracellular compartments. Still further fusion partners include affinity tags,
25 which facilitate purification of the polypeptide.

Fusion polypeptides may generally be prepared using standard techniques, including chemical conjugation. Preferably, a fusion polypeptide is expressed as a recombinant polypeptide, allowing the production of increased levels, relative to a non-fused polypeptide, in an expression system. Briefly, DNA sequences
30 encoding the polypeptide components may be assembled separately, and ligated into an appropriate expression vector. The 3' end of the DNA sequence encoding one

polypeptide component is ligated, with or without a peptide linker, to the 5' end of a DNA sequence encoding the second polypeptide component so that the reading frames of the sequences are in phase. This permits translation into a single fusion polypeptide that retains the biological activity of both component polypeptides.

- 5 A peptide linker sequence may be employed to separate the first and second polypeptide components by a distance sufficient to ensure that each polypeptide folds into its secondary and tertiary structures. Such a peptide linker sequence is incorporated into the fusion polypeptide using standard techniques well known in the art. Suitable peptide linker sequences may be chosen based on the following factors:
- 10 (1) their ability to adopt a flexible extended conformation; (2) their inability to adopt a secondary structure that could interact with functional epitopes on the first and second polypeptides; and (3) the lack of hydrophobic or charged residues that might react with the polypeptide functional epitopes. Preferred peptide linker sequences contain Gly, Asn and Ser residues. Other near neutral amino acids, such as Thr and Ala may also be
- 15 used in the linker sequence. Amino acid sequences which may be usefully employed as linkers include those disclosed in Maratea et al., *Gene* 40:39-46, 1985; Murphy et al., *Proc. Natl. Acad. Sci. USA* 83:8258-8262, 1986; U.S. Patent No. 4,935,233 and U.S. Patent No. 4,751,180. The linker sequence may generally be from 1 to about 50 amino acids in length. Linker sequences are not required when the first and second
- 20 polypeptides have non-essential N-terminal amino acid regions that can be used to separate the functional domains and prevent steric interference.

The ligated DNA sequences are operably linked to suitable transcriptional or translational regulatory elements. The regulatory elements responsible for expression of DNA are located only 5' to the DNA sequence encoding

25 the first polypeptides. Similarly, stop codons required to end translation and transcription termination signals are only present 3' to the DNA sequence encoding the second polypeptide.

The fusion polypeptide can comprise a polypeptide as described herein together with an unrelated immunogenic protein, such as an immunogenic protein

30 capable of eliciting a recall response. Examples of such proteins include tetanus,

tuberculosis and hepatitis proteins (*see, for example, Stoute et al. New Engl. J. Med., 336:86-91, 1997*).

In one preferred embodiment, the immunological fusion partner is derived from a *Mycobacterium* sp., such as a *Mycobacterium tuberculosis*-derived Ra12
5 fragment. Ra12 compositions and methods for their use in enhancing the expression and/or immunogenicity of heterologous polynucleotide/polypeptide sequences is described in U.S. Patent Application 60/158,585, the disclosure of which is incorporated herein by reference in its entirety. Briefly, Ra12 refers to a polynucleotide region that is a subsequence of a *Mycobacterium tuberculosis* MTB32A nucleic acid.
10 MTB32A is a serine protease of 32 KD molecular weight encoded by a gene in virulent and avirulent strains of *M. tuberculosis*. The nucleotide sequence and amino acid sequence of MTB32A have been described (for example, U.S. Patent Application 60/158,585; *see also, Skeiky et al., Infection and Immun. (1999) 67:3998-4007*, incorporated herein by reference). C-terminal fragments of the MTB32A coding
15 sequence express at high levels and remain as a soluble polypeptides throughout the purification process. Moreover, Ra12 may enhance the immunogenicity of heterologous immunogenic polypeptides with which it is fused. One preferred Ra12 fusion polypeptide comprises a 14 KD C-terminal fragment corresponding to amino acid residues 192 to 323 of MTB32A. Other preferred Ra12 polynucleotides generally
20 comprise at least about 15 consecutive nucleotides, at least about 30 nucleotides, at least about 60 nucleotides, at least about 100 nucleotides, at least about 200 nucleotides, or at least about 300 nucleotides that encode a portion of a Ra12 polypeptide. Ra12 polynucleotides may comprise a native sequence (*i.e.*, an endogenous sequence that encodes a Ra12 polypeptide or a portion thereof) or may comprise a variant of such a
25 sequence. Ra12 polynucleotide variants may contain one or more substitutions, additions, deletions and/or insertions such that the biological activity of the encoded fusion polypeptide is not substantially diminished, relative to a fusion polypeptide comprising a native Ra12 polypeptide. Variants preferably exhibit at least about 70% identity, more preferably at least about 80% identity and most preferably at least about
30 90% identity to a polynucleotide sequence that encodes a native Ra12 polypeptide or a portion thereof.

Within other preferred embodiments, an immunological fusion partner is derived from protein D, a surface protein of the gram-negative bacterium *Haemophilus influenza B* (WO 91/18926). Preferably, a protein D derivative comprises approximately the first third of the protein (e.g., the first N-terminal 100-110 amino acids), and a protein D derivative may be lipidated. Within certain preferred
5 embodiments, the first 109 residues of a Lipoprotein D fusion partner is included on the N-terminus to provide the polypeptide with additional exogenous T-cell epitopes and to increase the expression level in *E. coli* (thus functioning as an expression enhancer). The lipid tail ensures optimal presentation of the antigen to antigen presenting cells.

10 Other fusion partners include the non-structural protein from influenzae virus, NS1 (hemagglutinin). Typically, the N-terminal 81 amino acids are used, although different fragments that include T-helper epitopes may be used.

In another embodiment, the immunological fusion partner is the protein known as LYTA, or a portion thereof (preferably a C-terminal portion). LYTA is
15 derived from *Streptococcus pneumoniae*, which synthesizes an N-acetyl-L-alanine amidase known as amidase LYTA (encoded by the *LytA* gene; *Gene* 43:265-292, 1986). LYTA is an autolysin that specifically degrades certain bonds in the peptidoglycan backbone. The C-terminal domain of the LYTA protein is responsible for the affinity to the choline or to some choline analogues such as DEAE. This property has been
20 exploited for the development of *E. coli* C-LYTA expressing plasmids useful for expression of fusion proteins. Purification of hybrid proteins containing the C-LYTA fragment at the amino terminus has been described (see *Biotechnology* 10:795-798, 1992). Within a preferred embodiment, a repeat portion of LYTA may be incorporated into a fusion polypeptide. A repeat portion is found in the C-terminal region starting at
25 residue 178. A particularly preferred repeat portion incorporates residues 188-305.

Yet another illustrative embodiment involves fusion polypeptides, and the polynucleotides encoding them, wherein the fusion partner comprises a targeting signal capable of directing a polypeptide to the endosomal/lysosomal compartment, as described in U.S. Patent No. 5,633,234. An immunogenic polypeptide of the invention,
30 when fused with this targeting signal, will associate more efficiently with MHC class II

molecules and thereby provide enhanced in vivo stimulation of CD4⁺ T-cells specific for the polypeptide.

Polypeptides of the invention are prepared using any of a variety of well known synthetic and/or recombinant techniques, the latter of which are further described below. Polypeptides, portions and other variants generally less than about 150 amino acids can be generated by synthetic means, using techniques well known to those of ordinary skill in the art. In one illustrative example, such polypeptides are synthesized using any of the commercially available solid-phase techniques, such as the Merrifield solid-phase synthesis method, where amino acids are sequentially added to a growing amino acid chain. See Merrifield, *J. Am. Chem. Soc.* 85:2149-2146, 1963. Equipment for automated synthesis of polypeptides is commercially available from suppliers such as Perkin Elmer/Applied BioSystems Division (Foster City, CA), and may be operated according to the manufacturer's instructions.

In general, polypeptide compositions (including fusion polypeptides) of the invention are isolated. An "isolated" polypeptide is one that is removed from its original environment. For example, a naturally-occurring protein or polypeptide is isolated if it is separated from some or all of the coexisting materials in the natural system. Preferably, such polypeptides are also purified, *e.g.*, are at least about 90% pure, more preferably at least about 95% pure and most preferably at least about 99% pure.

Polynucleotide Compositions

The present invention, in other aspects, provides polynucleotide compositions. The terms "DNA" and "polynucleotide" are used essentially interchangeably herein to refer to a DNA molecule that has been isolated free of total genomic DNA of a particular species. "Isolated," as used herein, means that a polynucleotide is substantially away from other coding sequences, and that the DNA molecule does not contain large portions of unrelated coding DNA, such as large chromosomal fragments or other functional genes or polypeptide coding regions. Of course, this refers to the DNA molecule as originally isolated, and does not exclude genes or coding regions later added to the segment by the hand of man.

As will be understood by those skilled in the art, the polynucleotide compositions of this invention can include genomic sequences, extra-genomic and plasmid-encoded sequences and smaller engineered gene segments that express, or may be adapted to express, proteins, polypeptides, peptides and the like. Such segments may
5 be naturally isolated, or modified synthetically by the hand of man.

As will be also recognized by the skilled artisan, polynucleotides of the invention may be single-stranded (coding or antisense) or double-stranded, and may be DNA (genomic, cDNA or synthetic) or RNA molecules. RNA molecules may include HnRNA molecules, which contain introns and correspond to a DNA molecule in a one-
10 to-one manner, and mRNA molecules, which do not contain introns. Additional coding or non-coding sequences may, but need not, be present within a polynucleotide of the present invention, and a polynucleotide may, but need not, be linked to other molecules and/or support materials.

Polynucleotides may comprise a native sequence (*i.e.*, an endogenous
15 sequence that encodes a polypeptide/protein of the invention or a portion thereof) or may comprise a sequence that encodes a variant or derivative, preferably and immunogenic variant or derivative, of such a sequence.

Therefore, according to another aspect of the present invention, polynucleotide compositions are provided that comprise some or all of a polynucleotide
20 sequence set forth in any one of SEQ ID NO: 1-10,912, complements of a polynucleotide sequence set forth in any one of SEQ ID NO: 1-10,912, and degenerate variants of a polynucleotide sequence set forth in any one of SEQ ID NO: 1-10,912. In certain preferred embodiments, the polynucleotide sequences set forth herein encode immunogenic polypeptides, as described above.

25 In other related embodiments, the present invention provides polynucleotide variants having substantial identity to the sequences disclosed herein in SEQ ID NO: 1-10,912, for example those comprising at least 70% sequence identity, preferably at least 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% or higher, sequence identity compared to a polynucleotide sequence of this invention using the
30 methods described herein, (*e.g.*, BLAST analysis using standard parameters, as described below). One skilled in this art will recognize that these values can be

appropriately adjusted to determine corresponding identity of proteins encoded by two nucleotide sequences by taking into account codon degeneracy, amino acid similarity, reading frame positioning and the like.

Typically, polynucleotide variants will contain one or more substitutions, additions, deletions and/or insertions, preferably such that the immunogenicity of the polypeptide encoded by the variant polynucleotide is not substantially diminished relative to a polypeptide encoded by a polynucleotide sequence specifically set forth herein). The term "variants" should also be understood to encompass homologous genes of xenogenic origin.

10 In additional embodiments, the present invention provides polynucleotide fragments comprising or consisting of various lengths of contiguous stretches of sequence identical to or complementary to one or more of the sequences disclosed herein. For example, polynucleotides are provided by this invention that comprise or consist of at least about 10, 15, 20, 30, 40, 50, 75, 100, 150, 200, 300, 400,
15 500 or 1000 or more contiguous nucleotides of one or more of the sequences disclosed herein as well as all intermediate lengths there between. It will be readily understood that "intermediate lengths", in this context, means any length between the quoted values, such as 16, 17, 18, 19, *etc.*; 21, 22, 23, *etc.*; 30, 31, 32, *etc.*; 50, 51, 52, 53, *etc.*; 100, 101, 102, 103, *etc.*; 150, 151, 152, 153, *etc.*; including all integers through 200-
20 500; 500-1,000, and the like. A polynucleotide sequence as described here may be extended at one or both ends by additional nucleotides not found in the native sequence. This additional sequence may consist of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 nucleotides at either end of the disclosed sequence or at both ends of the disclosed sequence.

25 In another embodiment of the invention, polynucleotide compositions are provided that are capable of hybridizing under moderate to high stringency conditions to a polynucleotide sequence provided herein, or a fragment thereof, or a complementary sequence thereof. Hybridization techniques are well known in the art of molecular biology. For purposes of illustration, suitable moderately stringent conditions for
30 testing the hybridization of a polynucleotide of this invention with other polynucleotides include prewashing in a solution of 5 X SSC, 0.5% SDS, 1.0 mM EDTA (pH 8.0);

hybridizing at 50°C-60°C, 5 X SSC, overnight; followed by washing twice at 65°C for 20 minutes with each of 2X, 0.5X and 0.2X SSC containing 0.1% SDS. One skilled in the art will understand that the stringency of hybridization can be readily manipulated, such as by altering the salt content of the hybridization solution and/or the temperature at which the hybridization is performed. For example, in another embodiment, suitable highly stringent hybridization conditions include those described above, with the exception that the temperature of hybridization is increased, *e.g.*, to 60-65°C or 65-70°C.

In certain preferred embodiments, the polynucleotides described above, *e.g.*, polynucleotide variants, fragments and hybridizing sequences, encode polypeptides that are immunologically cross-reactive with a polypeptide sequence specifically set forth herein. In other preferred embodiments, such polynucleotides encode polypeptides that have a level of immunogenic activity of at least about 50%, preferably at least about 70%, and more preferably at least about 90% of that for a polypeptide sequence specifically set forth herein.

The polynucleotides of the present invention, or fragments thereof, regardless of the length of the coding sequence itself, may be combined with other DNA sequences, such as promoters, polyadenylation signals, additional restriction enzyme sites, multiple cloning sites, other coding segments, and the like, such that their overall length may vary considerably. It is therefore contemplated that a nucleic acid fragment of almost any length may be employed, with the total length preferably being limited by the ease of preparation and use in the intended recombinant DNA protocol. For example, illustrative polynucleotide segments with total lengths of about 10,000, about 5000, about 3000, about 2,000, about 1,000, about 500, about 200, about 100, about 50 base pairs in length, and the like, (including all intermediate lengths) are contemplated to be useful in many implementations of this invention.

When comparing polynucleotide sequences, two sequences are said to be "identical" if the sequence of nucleotides in the two sequences is the same when aligned for maximum correspondence, as described below. Comparisons between two sequences are typically performed by comparing the sequences over a comparison window to identify and compare local regions of sequence similarity. A "comparison

window" as used herein, refers to a segment of at least about 20 contiguous positions, usually 30 to about 75, 40 to about 50, in which a sequence may be compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned.

- 5 Optimal alignment of sequences for comparison may be conducted using the Megalign program in the Lasergene suite of bioinformatics software (DNASTAR, Inc., Madison, WI), using default parameters. This program embodies several alignment schemes described in the following references: Dayhoff, M.O. (1978) A model of evolutionary change in proteins – Matrices for detecting distant relationships.
- 10 In Dayhoff, M.O. (ed.) Atlas of Protein Sequence and Structure, National Biomedical Research Foundation, Washington DC Vol. 5, Suppl. 3, pp. 345-358; Hein J. (1990) Unified Approach to Alignment and Phylogenies pp. 626-645 *Methods in Enzymology* vol. 183, Academic Press, Inc., San Diego, CA; Higgins, D.G. and Sharp, P.M. (1989) *CABIOS* 5:151-153; Myers, E.W. and Muller W. (1988) *CABIOS* 4:11-17; Robinson,
- 15 E.D. (1971) *Comb. Theor* 11:105; Santou, N. Nes, M. (1987) *Mol. Biol. Evol.* 4:406-425; Sneath, P.H.A. and Sokal, R.R. (1973) *Numerical Taxonomy – the Principles and Practice of Numerical Taxonomy*, Freeman Press, San Francisco, CA; Wilbur, W.J. and Lipman, D.J. (1983) *Proc. Natl. Acad. Sci. USA* 80:726-730.

- Alternatively, optimal alignment of sequences for comparison may be
- 20 conducted by the local identity algorithm of Smith and Waterman (1981) *Add. APL Math* 2:482, by the identity alignment algorithm of Needleman and Wunsch (1970) *J. Mol. Biol.* 48:443, by the search for similarity methods of Pearson and Lipman (1988) *Proc. Natl. Acad. Sci. USA* 85: 2444, by computerized implementations of these algorithms (GAP, BESTFIT, BLAST, FASTA, and TFASTA in the Wisconsin Genetics
- 25 Software Package, Genetics Computer Group (GCG), 575 Science Dr., Madison, WI), or by inspection.

- One preferred example of algorithms that are suitable for determining percent sequence identity and sequence similarity are the BLAST and BLAST 2.0 algorithms, which are described in Altschul et al. (1977) *Nucl. Acids Res.* 25:3389-3402
- 30 and Altschul et al. (1990) *J. Mol. Biol.* 215:403-410, respectively. BLAST and BLAST 2.0 can be used, for example with the parameters described herein, to determine percent

sequence identity for the polynucleotides of the invention. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information. In one illustrative example, cumulative scores can be calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always >0) and N (penalty score for mismatching residues; always <0). Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, and expectation (E) of 10, and the BLOSUM62 scoring matrix (see Henikoff and Henikoff (1989) *Proc. Natl. Acad. Sci. USA* 89:10915) alignments, (B) of 50, expectation (E) of 10, M=5, N=-4 and a comparison of both strands.

Preferably, the "percentage of sequence identity" is determined by comparing two optimally aligned sequences over a window of comparison of at least 20 positions, wherein the portion of the polynucleotide sequence in the comparison window may comprise additions or deletions (*i.e.*, gaps) of 20 percent or less, usually 5 to 15 percent, or 10 to 12 percent, as compared to the reference sequences (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical nucleic acid bases occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the reference sequence (*i.e.*, the window size) and multiplying the results by 100 to yield the percentage of sequence identity.

It will be appreciated by those of ordinary skill in the art that, as a result of the degeneracy of the genetic code, there are many nucleotide sequences that encode a polypeptide as described herein. Some of these polynucleotides bear minimal homology to the nucleotide sequence of any native gene. Nonetheless, polynucleotides that vary due to differences in codon usage are specifically contemplated by the present invention. Further, alleles of the genes comprising the polynucleotide sequences

provided herein are within the scope of the present invention. Alleles are endogenous genes that are altered as a result of one or more mutations, such as deletions, additions and/or substitutions of nucleotides. The resulting mRNA and protein may, but need not, have an altered structure or function. Alleles may be identified using standard
5 techniques (such as hybridization, amplification and/or database sequence comparison).

Therefore, in another embodiment of the invention, a mutagenesis approach, such as site-specific mutagenesis, is employed for the preparation of immunogenic variants and/or derivatives of the polypeptides described herein. By this approach, specific modifications in a polypeptide sequence can be made through
10 mutagenesis of the underlying polynucleotides that encode them. These techniques provides a straightforward approach to prepare and test sequence variants, for example, incorporating one or more of the foregoing considerations, by introducing one or more nucleotide sequence changes into the polynucleotide.

Site-specific mutagenesis allows the production of mutants through the
15 use of specific oligonucleotide sequences which encode the DNA sequence of the desired mutation, as well as a sufficient number of adjacent nucleotides, to provide a primer sequence of sufficient size and sequence complexity to form a stable duplex on both sides of the deletion junction being traversed. Mutations may be employed in a selected polynucleotide sequence to improve, alter, decrease, modify, or otherwise
20 change the properties of the polynucleotide itself, and/or alter the properties, activity, composition, stability, or primary sequence of the encoded polypeptide.

In certain embodiments of the present invention, the inventors contemplate the mutagenesis of the disclosed polynucleotide sequences to alter one or more properties of the encoded polypeptide, such as the immunogenicity of a
25 polypeptide vaccine. The techniques of site-specific mutagenesis are well-known in the art, and are widely used to create variants of both polypeptides and polynucleotides. For example, site-specific mutagenesis is often used to alter a specific portion of a DNA molecule. In such embodiments, a primer comprising typically about 14 to about 25 nucleotides or so in length is employed, with about 5 to about 10 residues on both sides
30 of the junction of the sequence being altered.

As will be appreciated by those of skill in the art, site-specific mutagenesis techniques have often employed a phage vector that exists in both a single stranded and double stranded form. Typical vectors useful in site-directed mutagenesis include vectors such as the M13 phage. These phage are readily commercially-available and their use is generally well-known to those skilled in the art. Double-stranded plasmids are also routinely employed in site directed mutagenesis that eliminates the step of transferring the gene of interest from a plasmid to a phage.

In general, site-directed mutagenesis in accordance herewith is performed by first obtaining a single-stranded vector or melting apart of two strands of a double-stranded vector that includes within its sequence a DNA sequence that encodes the desired peptide. An oligonucleotide primer bearing the desired mutated sequence is prepared, generally synthetically. This primer is then annealed with the single-stranded vector, and subjected to DNA polymerizing enzymes such as *E. coli* polymerase I Klenow fragment, in order to complete the synthesis of the mutation-bearing strand. Thus, a heteroduplex is formed wherein one strand encodes the original non-mutated sequence and the second strand bears the desired mutation. This heteroduplex vector is then used to transform appropriate cells, such as *E. coli* cells, and clones are selected which include recombinant vectors bearing the mutated sequence arrangement.

The preparation of sequence variants of the selected peptide-encoding DNA segments using site-directed mutagenesis provides a means of producing potentially useful species and is not meant to be limiting as there are other ways in which sequence variants of peptides and the DNA sequences encoding them may be obtained. For example, recombinant vectors encoding the desired peptide sequence may be treated with mutagenic agents, such as hydroxylamine, to obtain sequence variants. Specific details regarding these methods and protocols are found in the teachings of Maloy *et al.*, 1994; Segal, 1976; Prokop and Bajpai, 1991; Kuby, 1994; and Maniatis *et al.*, 1982, each incorporated herein by reference, for that purpose.

As used herein, the term "oligonucleotide directed mutagenesis procedure" refers to template-dependent processes and vector-mediated propagation which result in an increase in the concentration of a specific nucleic acid molecule relative to its initial concentration, or in an increase in the concentration of a detectable

signal, such as amplification. As used herein, the term "oligonucleotide directed mutagenesis procedure" is intended to refer to a process that involves the template-dependent extension of a primer molecule. The term template dependent process refers to nucleic acid synthesis of an RNA or a DNA molecule wherein the sequence of the newly synthesized strand of nucleic acid is dictated by the well-known rules of complementary base pairing (see, for example, Watson, 1987). Typically, vector mediated methodologies involve the introduction of the nucleic acid fragment into a DNA or RNA vector, the clonal amplification of the vector, and the recovery of the amplified nucleic acid fragment. Examples of such methodologies are provided by U. S. Patent No. 4,237,224, specifically incorporated herein by reference in its entirety.

In another approach for the production of polypeptide variants of the present invention, recursive sequence recombination, as described in U.S. Patent No. 5,837,458, may be employed. In this approach, iterative cycles of recombination and screening or selection are performed to "evolve" individual polynucleotide variants of the invention having, for example, enhanced immunogenic activity.

In other embodiments of the present invention, the polynucleotide sequences provided herein can be advantageously used as probes or primers for nucleic acid hybridization. As such, it is contemplated that nucleic acid segments that comprise or consist of a sequence region of at least about a 15 nucleotide long contiguous sequence that has the same sequence as, or is complementary to, a 15 nucleotide long contiguous sequence disclosed herein will find particular utility. Longer contiguous identical or complementary sequences, *e.g.*, those of about 20, 30, 40, 50, 100, 200, 500, 1000 (including all intermediate lengths) and even up to full length sequences will also be of use in certain embodiments.

The ability of such nucleic acid probes to specifically hybridize to a sequence of interest will enable them to be of use in detecting the presence of complementary sequences in a given sample. However, other uses are also envisioned, such as the use of the sequence information for the preparation of mutant species primers, or primers for use in preparing other genetic constructions.

Polynucleotide molecules having sequence regions consisting of contiguous nucleotide stretches of 10-14, 15-20, 30, 50, or even of 100-200 nucleotides

or so (including intermediate lengths as well), identical or complementary to a polynucleotide sequence disclosed herein, are particularly contemplated as hybridization probes for use in, *e.g.*, Southern and Northern blotting. This would allow a gene product, or fragment thereof, to be analyzed, both in diverse cell types and also in various bacterial cells. The total size of fragment, as well as the size of the complementary stretch(es), will ultimately depend on the intended use or application of the particular nucleic acid segment. Smaller fragments will generally find use in hybridization embodiments, wherein the length of the contiguous complementary region may be varied, such as between about 15 and about 100 nucleotides, but larger contiguous complementarity stretches may be used, according to the length complementary sequences one wishes to detect.

The use of a hybridization probe of about 15-25 nucleotides in length allows the formation of a duplex molecule that is both stable and selective. Molecules having contiguous complementary sequences over stretches greater than 15 bases in length are generally preferred, though, in order to increase stability and selectivity of the hybrid, and thereby improve the quality and degree of specific hybrid molecules obtained. One will generally prefer to design nucleic acid molecules having gene-complementary stretches of 15 to 25 contiguous nucleotides, or even longer where desired.

Hybridization probes may be selected from any portion of any of the sequences disclosed herein. All that is required is to review the sequences set forth herein, or to any continuous portion of the sequences, from about 15-25 nucleotides in length up to and including the full length sequence, that one wishes to utilize as a probe or primer. The choice of probe and primer sequences may be governed by various factors. For example, one may wish to employ primers from towards the termini of the total sequence.

Small polynucleotide segments or fragments may be readily prepared by, for example, directly synthesizing the fragment by chemical means, as is commonly practiced using an automated oligonucleotide synthesizer. Also, fragments may be obtained by application of nucleic acid reproduction technology, such as the PCR™ technology of U. S. Patent 4,683,202 (incorporated herein by reference), by introducing

selected sequences into recombinant vectors for recombinant production, and by other recombinant DNA techniques generally known to those of skill in the art of molecular biology.

The nucleotide sequences of the invention may be used for their ability to
5 selectively form duplex molecules with complementary stretches of the entire gene or gene fragments of interest. Depending on the application envisioned, one will typically desire to employ varying conditions of hybridization to achieve varying degrees of selectivity of probe towards target sequence. For applications requiring high selectivity, one will typically desire to employ relatively stringent conditions to form the hybrids,
10 *e.g.*, one will select relatively low salt and/or high temperature conditions, such as provided by a salt concentration of from about 0.02 M to about 0.15 M salt at temperatures of from about 50°C to about 70°C. Such selective conditions tolerate little, if any, mismatch between the probe and the template or target strand, and would be particularly suitable for isolating related sequences.

15 Of course, for some applications, for example, where one desires to prepare mutants employing a mutant primer strand hybridized to an underlying template, less stringent (reduced stringency) hybridization conditions will typically be needed in order to allow formation of the heteroduplex. In these circumstances, one may desire to employ salt conditions such as those of from about 0.15 M to about 0.9 M
20 salt, at temperatures ranging from about 20°C to about 55°C. Cross-hybridizing species can thereby be readily identified as positively hybridizing signals with respect to control hybridizations. In any case, it is generally appreciated that conditions can be rendered more stringent by the addition of increasing amounts of formamide, which serves to destabilize the hybrid duplex in the same manner as increased temperature. Thus,
25 hybridization conditions can be readily manipulated, and thus will generally be a method of choice depending on the desired results.

According to another embodiment of the present invention, polynucleotide compositions comprising antisense oligonucleotides are provided. Antisense oligonucleotides have been demonstrated to be effective and targeted
30 inhibitors of protein synthesis, and, consequently, provide a therapeutic approach by which a disease can be treated by inhibiting the synthesis of proteins that contribute to

the disease. The efficacy of antisense oligonucleotides for inhibiting protein synthesis is well established. For example, the synthesis of polygalacturonase and the muscarine type 2 acetylcholine receptor are inhibited by antisense oligonucleotides directed to their respective mRNA sequences (U. S. Patent 5,739,119 and U. S. Patent 5,759,829).

5 Further, examples of antisense inhibition have been demonstrated with the nuclear protein cyclin, the multiple drug resistance gene (MDG1), ICAM-1, E-selectin, STK-1, striatal GABA_A receptor and human EGF (Jaskulski *et al.*, Science. 1988 Jun 10;240(4858):1544-6; Vasanthakumar and Ahmed, Cancer Commun. 1989;1(4):225-32; Peris *et al.*, Brain Res Mol Brain Res. 1998 Jun 15;57(2):310-20; U. S. Patent 10 5,801,154; U.S. Patent 5,789,573; U. S. Patent 5,718,709 and U.S. Patent 5,610,288). Antisense constructs have also been described that inhibit and can be used to treat a variety of abnormal cellular proliferations, *e.g.* cancer (U. S. Patent 5,747,470; U. S. Patent 5,591,317 and U. S. Patent 5,783,683).

Therefore, in certain embodiments, the present invention provides

15 oligonucleotide sequences that comprise all, or a portion of, any sequence that is capable of specifically binding to polynucleotide sequence described herein, or a complement thereof. In one embodiment, the antisense oligonucleotides comprise DNA or derivatives thereof. In another embodiment, the oligonucleotides comprise RNA or derivatives thereof. In a third embodiment, the oligonucleotides are modified DNAs

20 comprising a phosphorothioated modified backbone. In a fourth embodiment, the oligonucleotide sequences comprise peptide nucleic acids or derivatives thereof. In each case, preferred compositions comprise a sequence region that is complementary, and more preferably substantially-complementary, and even more preferably, completely complementary to one or more portions of polynucleotides disclosed herein.

25 Selection of antisense compositions specific for a given gene sequence is based upon analysis of the chosen target sequence and determination of secondary structure, T_m , binding energy, and relative stability. Antisense compositions may be selected based upon their relative inability to form dimers, hairpins, or other secondary structures that would reduce or prohibit specific binding to the target mRNA in a host cell. Highly

30 preferred target regions of the mRNA, are those which are at or near the AUG translation initiation codon, and those sequences which are substantially complementary

to 5' regions of the mRNA. These secondary structure analyses and target site selection considerations can be performed, for example, using v.4 of the OLIGO primer analysis software and/or the BLASTN 2.0.5 algorithm software (Altschul *et al.*, Nucleic Acids Res. 1997, 25(17):3389-402).

5 The use of an antisense delivery method employing a short peptide vector, termed MPG (27 residues), is also contemplated. The MPG peptide contains a hydrophobic domain derived from the fusion sequence of HIV gp41 and a hydrophilic domain from the nuclear localization sequence of SV40 T-antigen (Morris *et al.*, Nucleic Acids Res. 1997 Jul 15;25(14):2730-6). It has been demonstrated that several
10 molecules of the MPG peptide coat the antisense oligonucleotides and can be delivered into cultured mammalian cells in less than 1 hour with relatively high efficiency (90%). Further, the interaction with MPG strongly increases both the stability of the oligonucleotide to nuclease and the ability to cross the plasma membrane.

 According to another embodiment of the invention, the polynucleotide
15 compositions described herein are used in the design and preparation of ribozyme molecules for inhibiting expression of the tumor polypeptides and proteins of the present invention in tumor cells. Ribozymes are RNA-protein complexes that cleave nucleic acids in a site-specific fashion. Ribozymes have specific catalytic domains that possess endonuclease activity (Kim and Cech, Proc Natl Acad Sci U S A. 1987
20 Dec;84(24):8788-92; Forster and Symons, Cell. 1987 Apr 24;49(2):211-20). For example, a large number of ribozymes accelerate phosphoester transfer reactions with a high degree of specificity, often cleaving only one of several phosphoesters in an oligonucleotide substrate (Cech *et al.*, Cell. 1981 Dec;27(3 Pt 2):487-96; Michel and Westhof, J Mol Biol. 1990 Dec 5;216(3):585-610; Reinhold-Hurek and Shub, Nature.
25 1992 May 14;357(6374):173-6). This specificity has been attributed to the requirement that the substrate bind via specific base-pairing interactions to the internal guide sequence ("IGS") of the ribozyme prior to chemical reaction.

 Six basic varieties of naturally-occurring enzymatic RNAs are known presently. Each can catalyze the hydrolysis of RNA phosphodiester bonds *in trans* (and
30 thus can cleave other RNA molecules) under physiological conditions. In general, enzymatic nucleic acids act by first binding to a target RNA. Such binding occurs

through the target binding portion of a enzymatic nucleic acid which is held in close proximity to an enzymatic portion of the molecule that acts to cleave the target RNA. Thus, the enzymatic nucleic acid first recognizes and then binds a target RNA through complementary base-pairing, and once bound to the correct site, acts enzymatically to cut the target RNA. Strategic cleavage of such a target RNA will destroy its ability to direct synthesis of an encoded protein. After an enzymatic nucleic acid has bound and cleaved its RNA target, it is released from that RNA to search for another target and can repeatedly bind and cleave new targets.

The enzymatic nature of a ribozyme is advantageous over many technologies, such as antisense technology (where a nucleic acid molecule simply binds to a nucleic acid target to block its translation) since the concentration of ribozyme necessary to affect a therapeutic treatment is lower than that of an antisense oligonucleotide. This advantage reflects the ability of the ribozyme to act enzymatically. Thus, a single ribozyme molecule is able to cleave many molecules of target RNA. In addition, the ribozyme is a highly specific inhibitor, with the specificity of inhibition depending not only on the base pairing mechanism of binding to the target RNA, but also on the mechanism of target RNA cleavage. Single mismatches, or base-substitutions, near the site of cleavage can completely eliminate catalytic activity of a ribozyme. Similar mismatches in antisense molecules do not prevent their action (Woolf *et al.*, Proc Natl Acad Sci U S A. 1992 Aug 15;89(16):7305-9). Thus, the specificity of action of a ribozyme is greater than that of an antisense oligonucleotide binding the same RNA site.

The enzymatic nucleic acid molecule may be formed in a hammerhead, hairpin, a hepatitis δ virus, group I intron or RNaseP RNA (in association with an RNA guide sequence) or Neurospora VS RNA motif. Examples of hammerhead motifs are described by Rossi *et al.* Nucleic Acids Res. 1992 Sep 11;20(17):4559-65. Examples of hairpin motifs are described by Hampel *et al.* (Eur. Pat. Appl. Publ. No. EP 0360257), Hampel and Tritz, Biochemistry 1989 Jun 13;28(12):4929-33; Hampel *et al.*, Nucleic Acids Res. 1990 Jan 25;18(2):299-304 and U. S. Patent 5,631,359. An example of the hepatitis δ virus motif is described by Perrotta and Been, Biochemistry. 1992 Dec 1;31(47):11843-52; an example of the RNaseP motif is described by Guerrier-Takada

et al., Cell. 1983 Dec;35(3 Pt 2):849-57; Neurospora VS RNA ribozyme motif is described by Collins (Saville and Collins, Cell. 1990 May 18;61(4):685-96; Saville and Collins, Proc Natl Acad Sci U S A. 1991 Oct 1;88(19):8826-30; Collins and Olive, Biochemistry. 1993 Mar 23;32(11):2795-9); and an example of the Group I intron is described in (U. S. Patent 4,987,071). All that is important in an enzymatic nucleic acid molecule of this invention is that it has a specific substrate binding site which is complementary to one or more of the target gene RNA regions, and that it have nucleotide sequences within or surrounding that substrate binding site which impart an RNA cleaving activity to the molecule. Thus the ribozyme constructs need not be limited to specific motifs mentioned herein.

Ribozymes may be designed as described in Int. Pat. Appl. Publ. No. WO 93/23569 and Int. Pat. Appl. Publ. No. WO 94/02595, each specifically incorporated herein by reference) and synthesized to be tested *in vitro* and *in vivo*, as described. Such ribozymes can also be optimized for delivery. While specific examples are provided, those in the art will recognize that equivalent RNA targets in other species can be utilized when necessary.

Ribozyme activity can be optimized by altering the length of the ribozyme binding arms, or chemically synthesizing ribozymes with modifications that prevent their degradation by serum ribonucleases (see *e.g.*, Int. Pat. Appl. Publ. No. WO 92/07065; Int. Pat. Appl. Publ. No. WO 93/15187; Int. Pat. Appl. Publ. No. WO 91/03162; Eur. Pat. Appl. Publ. No. 92110298.4; U. S. Patent 5,334,711; and Int. Pat. Appl. Publ. No. WO 94/13688, which describe various chemical modifications that can be made to the sugar moieties of enzymatic RNA molecules), modifications which enhance their efficacy in cells, and removal of stem II bases to shorten RNA synthesis times and reduce chemical requirements.

Sullivan *et al.* (Int. Pat. Appl. Publ. No. WO 94/02595) describes the general methods for delivery of enzymatic RNA molecules. Ribozymes may be administered to cells by a variety of methods known to those familiar to the art, including, but not restricted to, encapsulation in liposomes, by iontophoresis, or by incorporation into other vehicles, such as hydrogels, cyclodextrins, biodegradable nanocapsules, and bioadhesive microspheres. For some indications, ribozymes may be

directly delivered *ex vivo* to cells or tissues with or without the aforementioned vehicles. Alternatively, the RNA/vehicle combination may be locally delivered by direct inhalation, by direct injection or by use of a catheter, infusion pump or stent. Other routes of delivery include, but are not limited to, intravascular, intramuscular, subcutaneous or joint injection, aerosol inhalation, oral (tablet or pill form), topical, systemic, ocular, intraperitoneal and/or intrathecal delivery. More detailed descriptions of ribozyme delivery and administration are provided in Int. Pat. Appl. Publ. No. WO 94/02595 and Int. Pat. Appl. Publ. No. WO 93/23569, each specifically incorporated herein by reference.

Another means of accumulating high concentrations of a ribozyme(s) within cells is to incorporate the ribozyme-encoding sequences into a DNA expression vector. Transcription of the ribozyme sequences are driven from a promoter for eukaryotic RNA polymerase I (pol I), RNA polymerase II (pol II), or RNA polymerase III (pol III). Transcripts from pol II or pol III promoters will be expressed at high levels in all cells; the levels of a given pol II promoter in a given cell type will depend on the nature of the gene regulatory sequences (enhancers, silencers, *etc.*) present nearby. Prokaryotic RNA polymerase promoters may also be used, providing that the prokaryotic RNA polymerase enzyme is expressed in the appropriate cells. Ribozymes expressed from such promoters have been shown to function in mammalian cells. Such transcription units can be incorporated into a variety of vectors for introduction into mammalian cells, including but not restricted to, plasmid DNA vectors, viral DNA vectors (such as adenovirus or adeno-associated vectors), or viral RNA vectors (such as retroviral, semliki forest virus, sindbis virus vectors).

In another embodiment of the invention, peptide nucleic acids (PNAs) compositions are provided. PNA is a DNA mimic in which the nucleobases are attached to a pseudopeptide backbone (Good and Nielsen, *Antisense Nucleic Acid Drug Dev.* 1997 7(4) 431-37). PNA is able to be utilized in a number methods that traditionally have used RNA or DNA. Often PNA sequences perform better in techniques than the corresponding RNA or DNA sequences and have utilities that are not inherent to RNA or DNA. A review of PNA including methods of making, characteristics of, and methods of using, is provided by Corey (*Trends Biotechnol* 1997

Jun;15(6):224-9). As such, in certain embodiments, one may prepare PNA sequences that are complementary to one or more portions of the ACE mRNA sequence, and such PNA compositions may be used to regulate, alter, decrease, or reduce the translation of ACE-specific mRNA, and thereby alter the level of ACE activity in a host cell to which
5 such PNA compositions have been administered.

PNAs have 2-aminoethyl-glycine linkages replacing the normal phosphodiester backbone of DNA (Nielsen *et al.*, *Science* 1991 Dec 6;254(5037):1497-500; Hanvey *et al.*, *Science*. 1992 Nov 27;258(5087):1481-5; Hyrup and Nielsen, *Bioorg Med Chem*. 1996 Jan;4(1):5-23). This chemistry has three important
10 consequences: firstly, in contrast to DNA or phosphorothioate oligonucleotides, PNAs are neutral molecules; secondly, PNAs are achiral, which avoids the need to develop a stereoselective synthesis; and thirdly, PNA synthesis uses standard Boc or Fmoc protocols for solid-phase peptide synthesis, although other methods, including a modified Merrifield method, have been used.

15 PNA monomers or ready-made oligomers are commercially available from PerSeptive Biosystems (Framingham, MA). PNA syntheses by either Boc or Fmoc protocols are straightforward using manual or automated protocols (Norton *et al.*, *Bioorg Med Chem*. 1995 Apr;3(4):437-45). The manual protocol lends itself to the production of chemically modified PNAs or the simultaneous synthesis of families of
20 closely related PNAs.

As with peptide synthesis, the success of a particular PNA synthesis will depend on the properties of the chosen sequence. For example, while in theory PNAs can incorporate any combination of nucleotide bases, the presence of adjacent purines can lead to deletions of one or more residues in the product. In expectation of this
25 difficulty, it is suggested that, in producing PNAs with adjacent purines, one should repeat the coupling of residues likely to be added inefficiently. This should be followed by the purification of PNAs by reverse-phase high-pressure liquid chromatography, providing yields and purity of product similar to those observed during the synthesis of peptides.

30 Modifications of PNAs for a given application may be accomplished by coupling amino acids during solid-phase synthesis or by attaching compounds that

contain a carboxylic acid group to the exposed N-terminal amine. Alternatively, PNAs can be modified after synthesis by coupling to an introduced lysine or cysteine. The ease with which PNAs can be modified facilitates optimization for better solubility or for specific functional requirements. Once synthesized, the identity of PNAs and their derivatives can be confirmed by mass spectrometry. Several studies have made and utilized modifications of PNAs (for example, Norton *et al.*, Bioorg Med Chem. 1995 Apr;3(4):437-45; Petersen *et al.*, J Pept Sci. 1995 May-Jun;1(3):175-83; Orum *et al.*, Biotechniques. 1995 Sep;19(3):472-80; Footer *et al.*, Biochemistry. 1996 Aug 20;35(33):10673-9; Griffith *et al.*, Nucleic Acids Res. 1995 Aug 11;23(15):3003-8; Pardridge *et al.*, Proc Natl Acad Sci U S A. 1995 Jun 6;92(12):5592-6; Boffa *et al.*, Proc Natl Acad Sci U S A. 1995 Mar 14;92(6):1901-5; Gambacorti-Passerini *et al.*, Blood. 1996 Aug 15;88(4):1411-7; Armitage *et al.*, Proc Natl Acad Sci U S A. 1997 Nov 11;94(23):12320-5; Seeger *et al.*, Biotechniques. 1997 Sep;23(3):512-7). U.S. Patent No. 5,700,922 discusses PNA-DNA-PNA chimeric molecules and their uses in diagnostics, modulating protein in organisms, and treatment of conditions susceptible to therapeutics.

Methods of characterizing the antisense binding properties of PNAs are discussed in Rose (Anal Chem. 1993 Dec 15;65(24):3545-9) and Jensen *et al.* (Biochemistry. 1997 Apr 22;36(16):5072-7). Rose uses capillary gel electrophoresis to determine binding of PNAs to their complementary oligonucleotide, measuring the relative binding kinetics and stoichiometry. Similar types of measurements were made by Jensen *et al.* using BIAcore™ technology.

Other applications of PNAs that have been described and will be apparent to the skilled artisan include use in DNA strand invasion, antisense inhibition, mutational analysis, enhancers of transcription, nucleic acid purification, isolation of transcriptionally active genes, blocking of transcription factor binding, genome cleavage, biosensors, *in situ* hybridization, and the like.

Polynucleotide Identification, Characterization and Expression

Polynucleotides compositions of the present invention may be identified, prepared and/or manipulated using any of a variety of well established techniques (see

generally, Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratories, Cold Spring Harbor, NY, 1989, and other like references). For example, a polynucleotide may be identified, as described in more detail below, by screening a microarray of cDNAs for tumor-associated expression (*i.e.*, expression that is at least two fold greater in a tumor than in normal tissue, as determined using a representative assay provided herein). Such screens may be performed, for example, using the microarray technology of Affymetrix, Inc. (Santa Clara, CA) according to the manufacturer's instructions (and essentially as described by Schena et al., *Proc. Natl. Acad. Sci. USA* 93:10614-10619, 1996 and Heller et al., *Proc. Natl. Acad. Sci. USA* 94:2150-2155, 1997). Alternatively, polynucleotides may be amplified from cDNA prepared from cells expressing the proteins described herein, such as tumor cells.

Many template dependent processes are available to amplify a target sequences of interest present in a sample. One of the best known amplification methods is the polymerase chain reaction (PCRTM) which is described in detail in U.S. Patent Nos. 4,683,195, 4,683,202 and 4,800,159, each of which is incorporated herein by reference in its entirety. Briefly, in PCRTM, two primer sequences are prepared which are complementary to regions on opposite complementary strands of the target sequence. An excess of deoxynucleoside triphosphates is added to a reaction mixture along with a DNA polymerase (*e.g.*, *Taq* polymerase). If the target sequence is present in a sample, the primers will bind to the target and the polymerase will cause the primers to be extended along the target sequence by adding on nucleotides. By raising and lowering the temperature of the reaction mixture, the extended primers will dissociate from the target to form reaction products, excess primers will bind to the target and to the reaction product and the process is repeated. Preferably reverse transcription and PCRTM amplification procedure may be performed in order to quantify the amount of mRNA amplified. Polymerase chain reaction methodologies are well known in the art.

Any of a number of other template dependent processes, many of which are variations of the PCRTM amplification technique, are readily known and available in the art. Illustratively, some such methods include the ligase chain reaction (referred to as LCR), described, for example, in Eur. Pat. Appl. Publ. No. 320,308 and U.S. Patent

No. 4,883,750; Qbeta Replicase, described in PCT Intl. Pat. Appl. Publ. No. PCT/US87/00880; Strand Displacement Amplification (SDA) and Repair Chain Reaction (RCR). Still other amplification methods are described in Great Britain Pat. Appl. No. 2 202 328, and in PCT Intl. Pat. Appl. Publ. No. PCT/US89/01025. Other
5 nucleic acid amplification procedures include transcription-based amplification systems (TAS) (PCT Intl. Pat. Appl. Publ. No. WO 88/10315), including nucleic acid sequence based amplification (NASBA) and 3SR. Eur. Pat. Appl. Publ. No. 329,822 describes a nucleic acid amplification process involving cyclically synthesizing single-stranded RNA ("ssRNA"), ssDNA, and double-stranded DNA (dsDNA). PCT Intl. Pat. Appl.
10 Publ. No. WO 89/06700 describes a nucleic acid sequence amplification scheme based on the hybridization of a promoter/primer sequence to a target single-stranded DNA ("ssDNA") followed by transcription of many RNA copies of the sequence. Other amplification methods such as "RACE" (Frohman, 1990), and "one-sided PCR" (Ohara, 1989) are also well-known to those of skill in the art.

15 An amplified portion of a polynucleotide of the present invention may be used to isolate a full length gene from a suitable library (e.g., a tumor cDNA library) using well known techniques. Within such techniques, a library (cDNA or genomic) is screened using one or more polynucleotide probes or primers suitable for amplification. Preferably, a library is size-selected to include larger molecules. Random primed
20 libraries may also be preferred for identifying 5' and upstream regions of genes. Genomic libraries are preferred for obtaining introns and extending 5' sequences.

For hybridization techniques, a partial sequence may be labeled (e.g., by nick-translation or end-labeling with ³²P) using well known techniques. A bacterial or bacteriophage library is then generally screened by hybridizing filters containing
25 denatured bacterial colonies (or lawns containing phage plaques) with the labeled probe (see Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratories, Cold Spring Harbor, NY, 1989). Hybridizing colonies or plaques are selected and expanded, and the DNA is isolated for further analysis. cDNA clones may be analyzed to determine the amount of additional sequence by, for example, PCR using
30 a primer from the partial sequence and a primer from the vector. Restriction maps and partial sequences may be generated to identify one or more overlapping clones. The

complete sequence may then be determined using standard techniques, which may involve generating a series of deletion clones. The resulting overlapping sequences can then be assembled into a single contiguous sequence. A full length cDNA molecule can be generated by ligating suitable fragments, using well known techniques.

5 Alternatively, amplification techniques, such as those described above, can be useful for obtaining a full length coding sequence from a partial cDNA sequence. One such amplification technique is inverse PCR (*see* Triglia et al., *Nucl. Acids Res.* 16:8186, 1988), which uses restriction enzymes to generate a fragment in the known region of the gene. The fragment is then circularized by intramolecular ligation and
10 used as a template for PCR with divergent primers derived from the known region. Within an alternative approach, sequences adjacent to a partial sequence may be retrieved by amplification with a primer to a linker sequence and a primer specific to a known region. The amplified sequences are typically subjected to a second round of amplification with the same linker primer and a second primer specific to the known
15 region. A variation on this procedure, which employs two primers that initiate extension in opposite directions from the known sequence, is described in WO 96/38591. Another such technique is known as "rapid amplification of cDNA ends" or RACE. This technique involves the use of an internal primer and an external primer, which hybridizes to a polyA region or vector sequence, to identify sequences that are 5'
20 and 3' of a known sequence. Additional techniques include capture PCR (Lagerstrom et al., *PCR Methods Applic.* 1:111-19, 1991) and walking PCR (Parker et al., *Nucl. Acids Res.* 19:3055-60, 1991). Other methods employing amplification may also be employed to obtain a full length cDNA sequence.

 In certain instances, it is possible to obtain a full length cDNA sequence
25 by analysis of sequences provided in an expressed sequence tag (EST) database, such as that available from GenBank. Searches for overlapping ESTs may generally be performed using well known programs (*e.g.*, NCBI BLAST searches), and such ESTs may be used to generate a contiguous full length sequence. Full length DNA sequences may also be obtained by analysis of genomic fragments.

30 In other embodiments of the invention, polynucleotide sequences or fragments thereof which encode polypeptides of the invention, or fusion proteins or

functional equivalents thereof, may be used in recombinant DNA molecules to direct expression of a polypeptide in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences that encode substantially the same or a functionally equivalent amino acid sequence may be produced and these sequences may
5 be used to clone and express a given polypeptide.

As will be understood by those of skill in the art, it may be advantageous in some instances to produce polypeptide-encoding nucleotide sequences possessing non-naturally occurring codons. For example, codons preferred by a particular prokaryotic or eukaryotic host can be selected to increase the rate of protein expression
10 or to produce a recombinant RNA transcript having desirable properties, such as a half-life which is longer than that of a transcript generated from the naturally occurring sequence.

Moreover, the polynucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter polypeptide
15 encoding sequences for a variety of reasons, including but not limited to, alterations which modify the cloning, processing, and/or expression of the gene product. For example, DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. In addition, site-directed mutagenesis may be used to insert new restriction
20 sites, alter glycosylation patterns, change codon preference, produce splice variants, or introduce mutations, and so forth.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences may be ligated to a heterologous sequence to encode a fusion protein. For example, to screen peptide libraries for inhibitors of
25 polypeptide activity, it may be useful to encode a chimeric protein that can be recognized by a commercially available antibody. A fusion protein may also be engineered to contain a cleavage site located between the polypeptide-encoding sequence and the heterologous protein sequence, so that the polypeptide may be cleaved and purified away from the heterologous moiety.

30 Sequences encoding a desired polypeptide may be synthesized, in whole or in part, using chemical methods well known in the art (see Caruthers, M. H. et al.

(1980) *Nucl. Acids Res. Symp. Ser.* 215-223, Horn, T. et al. (1980) *Nucl. Acids Res. Symp. Ser.* 225-232). Alternatively, the protein itself may be produced using chemical methods to synthesize the amino acid sequence of a polypeptide, or a portion thereof. For example, peptide synthesis can be performed using various solid-phase techniques
5 (Roberge, J. Y. et al. (1995) *Science* 269:202-204) and automated synthesis may be achieved, for example, using the ABI 431A Peptide Synthesizer (Perkin Elmer, Palo Alto, CA).

A newly synthesized peptide may be substantially purified by preparative high performance liquid chromatography (e.g., Creighton, T. (1983) *Proteins, Structures and Molecular Principles*, WH Freeman and Co., New York, N.Y.) or other comparable
10 techniques available in the art. The composition of the synthetic peptides may be confirmed by amino acid analysis or sequencing (e.g., the Edman degradation procedure). Additionally, the amino acid sequence of a polypeptide, or any part thereof, may be altered during direct synthesis and/or combined using chemical methods with
15 sequences from other proteins, or any part thereof, to produce a variant polypeptide.

In order to express a desired polypeptide, the nucleotide sequences encoding the polypeptide, or functional equivalents, may be inserted into appropriate expression vector, i.e., a vector which contains the necessary elements for the transcription and translation of the inserted coding sequence. Methods which are well
20 known to those skilled in the art may be used to construct expression vectors containing sequences encoding a polypeptide of interest and appropriate transcriptional and translational control elements. These methods include *in vitro* recombinant DNA techniques, synthetic techniques, and *in vivo* genetic recombination. Such techniques are described, for example, in Sambrook, J. et al. (1989) *Molecular Cloning, A*
25 *Laboratory Manual*, Cold Spring Harbor Press, Plainview, N.Y., and Ausubel, F. M. et al. (1989) *Current Protocols in Molecular Biology*, John Wiley & Sons, New York, N.Y.

A variety of expression vector/host systems may be utilized to contain and express polynucleotide sequences. These include, but are not limited to,
30 microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors;

insect cell systems infected with virus expression vectors (e.g., baculovirus); plant cell systems transformed with virus expression vectors (e.g., cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems.

5 The "control elements" or "regulatory sequences" present in an expression vector are those non-translated regions of the vector--enhancers, promoters, 5' and 3' untranslated regions--which interact with host cellular proteins to carry out transcription and translation. Such elements may vary in their strength and specificity. Depending on the vector system and host utilized, any number of suitable transcription
10 and translation elements, including constitutive and inducible promoters, may be used. For example, when cloning in bacterial systems, inducible promoters such as the hybrid lacZ promoter of the pBLUESCRIPT phagemid (Stratagene, La Jolla, Calif.) or pSPORT1 plasmid (Gibco BRL, Gaithersburg, MD) and the like may be used. In mammalian cell systems, promoters from mammalian genes or from mammalian viruses
15 are generally preferred. If it is necessary to generate a cell line that contains multiple copies of the sequence encoding a polypeptide, vectors based on SV40 or EBV may be advantageously used with an appropriate selectable marker.

 In bacterial systems, any of a number of expression vectors may be selected depending upon the use intended for the expressed polypeptide. For example,
20 when large quantities are needed, for example for the induction of antibodies, vectors which direct high level expression of fusion proteins that are readily purified may be used. Such vectors include, but are not limited to, the multifunctional *E. coli* cloning and expression vectors such as pBLUESCRIPT (Stratagene), in which the sequence encoding the polypeptide of interest may be ligated into the vector in frame with
25 sequences for the amino-terminal Met and the subsequent 7 residues of .beta.-galactosidase so that a hybrid protein is produced; pIN vectors (Van Heeke, G. and S. M. Schuster (1989) *J. Biol. Chem.* 264:5503-5509); and the like. pGEX Vectors (Promega, Madison, Wis.) may also be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are
30 soluble and can easily be purified from lysed cells by adsorption to glutathione-agarose beads followed by elution in the presence of free glutathione. Proteins made in such

systems may be designed to include heparin, thrombin, or factor XA protease cleavage sites so that the cloned polypeptide of interest can be released from the GST moiety at will.

In the yeast, *Saccharomyces cerevisiae*, a number of vectors containing
5 constitutive or inducible promoters such as alpha factor, alcohol oxidase, and PGH may be used. For reviews, see Ausubel et al. (supra) and Grant et al. (1987) *Methods Enzymol.* 153:516-544.

In cases where plant expression vectors are used, the expression of sequences encoding polypeptides may be driven by any of a number of promoters. For
10 example, viral promoters such as the 35S and 19S promoters of CaMV may be used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) *EMBO J.* 6:307-311. Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used (Coruzzi, G. et al. (1984) *EMBO J.* 3:1671-1680; Broglie, R. et al. (1984) *Science* 224:838-843; and Winter, J. et al. (1991)
15 *Results Probl. Cell Differ.* 17:85-105). These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. Such techniques are described in a number of generally available reviews (see, for example, Hobbs, S. or Murry, L. E. in McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York, N.Y.; pp. 191-196).

20 An insect system may also be used to express a polypeptide of interest. For example, in one such system, *Autographa californica* nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes in *Spodoptera frugiperda* cells or in *Trichoplusia* larvae. The sequences encoding the polypeptide may be cloned into a non-essential region of the virus, such as the polyhedrin gene, and placed under control
25 of the polyhedrin promoter. Successful insertion of the polypeptide-encoding sequence will render the polyhedrin gene inactive and produce recombinant virus lacking coat protein. The recombinant viruses may then be used to infect, for example, *S. frugiperda* cells or *Trichoplusia* larvae in which the polypeptide of interest may be expressed (Engelhard, E. K. et al. (1994) *Proc. Natl. Acad. Sci.* 91 :3224-3227).

30 In mammalian host cells, a number of viral-based expression systems are generally available. For example, in cases where an adenovirus is used as an expression

vector, sequences encoding a polypeptide of interest may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain a viable virus which is capable of expressing the polypeptide in infected host
5 cells (Logan, J. and Shenk, T. (1984) *Proc. Natl. Acad. Sci.* 81:3655-3659). In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells.

Specific initiation signals may also be used to achieve more efficient translation of sequences encoding a polypeptide of interest. Such signals include the
10 ATG initiation codon and adjacent sequences. In cases where sequences encoding the polypeptide, its initiation codon, and upstream sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a portion thereof, is inserted, exogenous translational control signals including the ATG initiation
15 codon should be provided. Furthermore, the initiation codon should be in the correct reading frame to ensure translation of the entire insert. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers which are appropriate for the particular cell system which is used, such as those described in the
20 literature (Scharf, D. et al. (1994) *Results Probl. Cell Differ.* 20:125-162).

In addition, a host cell strain may be chosen for its ability to modulate the expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation.
25 Post-translational processing which cleaves a "prepro" form of the protein may also be used to facilitate correct insertion, folding and/or function. Different host cells such as CHO, COS, HeLa, MDCK, HEK293, and WI38, which have specific cellular machinery and characteristic mechanisms for such post-translational activities, may be chosen to ensure the correct modification and processing of the foreign protein.

30 For long-term, high-yield production of recombinant proteins, stable expression is generally preferred. For example, cell lines which stably express a

polynucleotide of interest may be transformed using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for 1-2 days in an enriched media before
5 they are switched to selective media. The purpose of the selectable marker is to confer resistance to selection, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be proliferated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed
10 cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase (Wigler, M. et al. (1977) *Cell* 11:223-32) and adenine phosphoribosyltransferase (Lowy, I. et al. (1990) *Cell* 22:817-23) genes which can be employed in tk.sup.- or apt.sup.- cells, respectively. Also, antimetabolite, antibiotic or herbicide resistance can be used as the basis for selection; for example, dhfr which confers resistance to
15 methotrexate (Wigler, M. et al. (1980) *Proc. Natl. Acad. Sci.* 77:3567-70); npt, which confers resistance to the aminoglycosides, neomycin and G-418 (Colbere-Garapin, F. et al (1981) *J. Mol. Biol.* 150:1-14); and als or pat, which confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively (Murry, *supra*). Additional selectable genes have been described, for example, trpB, which allows cells
20 to utilize indole in place of tryptophan, or hisD, which allows cells to utilize histinol in place of histidine (Hartman, S. C. and R. C. Mulligan (1988) *Proc. Natl. Acad. Sci.* 85:8047-51). The use of visible markers has gained popularity with such markers as anthocyanins, beta-glucuronidase and its substrate GUS, and luciferase and its substrate luciferin, being widely used not only to identify transformants, but also to quantify the
25 amount of transient or stable protein expression attributable to a specific vector system (Rhodes, C. A. et al. (1995) *Methods Mol. Biol.* 55:121-131).

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, its presence and expression may need to be confirmed. For example, if the sequence encoding a polypeptide is inserted within a
30 marker gene sequence, recombinant cells containing sequences can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem

with a polypeptide-encoding sequence under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

Alternatively, host cells that contain and express a desired
5 polynucleotide sequence may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations and protein bioassay or immunoassay techniques which include, for example, membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein.

10 A variety of protocols for detecting and measuring the expression of polynucleotide-encoded products, using either polyclonal or monoclonal antibodies specific for the product are known in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal
15 antibodies reactive to two non-interfering epitopes on a given polypeptide may be preferred for some applications, but a competitive binding assay may also be employed. These and other assays are described, among other places, in Hampton, R. et al. (1990; Serological Methods, a Laboratory Manual, APS Press, St Paul, Minn.) and Maddox, D. E. et al. (1983; *J. Exp. Med.* 158:1211-1216).

20 A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides include oligolabeling, nick translation, end-labeling or PCR amplification using a labeled nucleotide. Alternatively, the sequences, or any portions
25 thereof may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits. Suitable reporter molecules or labels, which may be used
30 include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with a polynucleotide sequence of interest may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a recombinant cell may be secreted or contained intracellularly depending on the sequence and/or the vector used. As will be understood
5 by those of skill in the art, expression vectors containing polynucleotides of the invention may be designed to contain signal sequences which direct secretion of the encoded polypeptide through a prokaryotic or eukaryotic cell membrane. Other recombinant constructions may be used to join sequences encoding a polypeptide of interest to nucleotide sequence encoding a polypeptide domain which will facilitate
10 purification of soluble proteins. Such purification facilitating domains include, but are not limited to, metal chelating peptides such as histidine-tryptophan modules that allow purification on immobilized metals, protein A domains that allow purification on immobilized immunoglobulin, and the domain utilized in the FLAGS extension/affinity purification system (Immunex Corp., Seattle, Wash.). The inclusion of cleavable linker
15 sequences such as those specific for Factor XA or enterokinase (Invitrogen, San Diego, Calif.) between the purification domain and the encoded polypeptide may be used to facilitate purification. One such expression vector provides for expression of a fusion protein containing a polypeptide of interest and a nucleic acid encoding 6 histidine residues preceding a thioredoxin or an enterokinase cleavage site. The histidine residues
20 facilitate purification on IMIAC (immobilized metal ion affinity chromatography) as described in Porath, J. et al. (1992, *Prot. Exp. Purif.* 3:263-281) while the enterokinase cleavage site provides a means for purifying the desired polypeptide from the fusion protein. A discussion of vectors which contain fusion proteins is provided in Kroll, D. J. et al. (1993; *DNA Cell Biol.* 12:441-453).

25 In addition to recombinant production methods, polypeptides of the invention, and fragments thereof, may be produced by direct peptide synthesis using solid-phase techniques (Merrifield J. (1963) *J. Am. Chem. Soc.* 85:2149-2154). Protein synthesis may be performed using manual techniques or by automation. Automated synthesis may be achieved, for example, using Applied Biosystems 431A Peptide
30 Synthesizer (Perkin Elmer). Alternatively, various fragments may be chemically

synthesized separately and combined using chemical methods to produce the full length molecule.

Antibody Compositions, Fragments Thereof and Other Binding Agents

According to another aspect, the present invention further provides
5 binding agents, such as antibodies and antigen-binding fragments thereof, that exhibit immunological binding to a tumor polypeptide disclosed herein, or to a portion, variant or derivative thereof. An antibody, or antigen-binding fragment thereof, is said to "specifically bind," "immunologically bind," and/or is "immunologically reactive" to a polypeptide of the invention if it reacts at a detectable level (within, for example, an
10 ELISA assay) with the polypeptide, and does not react detectably with unrelated polypeptides under similar conditions.

Immunological binding, as used in this context, generally refers to the non-covalent interactions of the type which occur between an immunoglobulin molecule and an antigen for which the immunoglobulin is specific. The strength, or
15 affinity of immunological binding interactions can be expressed in terms of the dissociation constant (K_d) of the interaction, wherein a smaller K_d represents a greater affinity. Immunological binding properties of selected polypeptides can be quantified using methods well known in the art. One such method entails measuring the rates of antigen-binding site/antigen complex formation and dissociation, wherein those rates
20 depend on the concentrations of the complex partners, the affinity of the interaction, and on geometric parameters that equally influence the rate in both directions. Thus, both the "on rate constant" (K_{on}) and the "off rate constant" (K_{off}) can be determined by calculation of the concentrations and the actual rates of association and dissociation. The ratio of K_{off}/K_{on} enables cancellation of all parameters not related to affinity, and is
25 thus equal to the dissociation constant K_d . See, generally, Davies et al. (1990) Annual Rev. Biochem. 59:439-473.

An "antigen-binding site," or "binding portion" of an antibody refers to the part of the immunoglobulin molecule that participates in antigen binding. The antigen binding site is formed by amino acid residues of the N-terminal variable ("V")
30 regions of the heavy ("H") and light ("L") chains. Three highly divergent stretches

within the V regions of the heavy and light chains are referred to as "hypervariable regions" which are interposed between more conserved flanking stretches known as "framework regions," or "FRs". Thus the term "FR" refers to amino acid sequences which are naturally found between and adjacent to hypervariable regions in immunoglobulins. In an antibody molecule, the three hypervariable regions of a light chain and the three hypervariable regions of a heavy chain are disposed relative to each other in three dimensional space to form an antigen-binding surface. The antigen-binding surface is complementary to the three-dimensional surface of a bound antigen, and the three hypervariable regions of each of the heavy and light chains are referred to as "complementarity-determining regions," or "CDRs."

Binding agents may be further capable of differentiating between patients with and without a cancer, such as ovarian cancer, using the representative assays provided herein. For example, antibodies or other binding agents that bind to a tumor protein will preferably generate a signal indicating the presence of a cancer in at least about 20% of patients with the disease, more preferably at least about 30% of patients. Alternatively, or in addition, the antibody will generate a negative signal indicating the absence of the disease in at least about 90% of individuals without the cancer. To determine whether a binding agent satisfies this requirement, biological samples (*e.g.*, blood, sera, sputum, urine and/or tumor biopsies) from patients with and without a cancer (as determined using standard clinical tests) may be assayed as described herein for the presence of polypeptides that bind to the binding agent. Preferably, a statistically significant number of samples with and without the disease will be assayed. Each binding agent should satisfy the above criteria; however, those of ordinary skill in the art will recognize that binding agents may be used in combination to improve sensitivity.

Any agent that satisfies the above requirements may be a binding agent. For example, a binding agent may be a ribosome, with or without a peptide component, an RNA molecule or a polypeptide. In a preferred embodiment, a binding agent is an antibody or an antigen-binding fragment thereof. Antibodies may be prepared by any of a variety of techniques known to those of ordinary skill in the art. *See, e.g.*, Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988. In

general, antibodies can be produced by cell culture techniques, including the generation of monoclonal antibodies as described herein, or via transfection of antibody genes into suitable bacterial or mammalian cell hosts, in order to allow for the production of recombinant antibodies. In one technique, an immunogen comprising the polypeptide is initially injected into any of a wide variety of mammals (*e.g.*, mice, rats, rabbits, sheep or goats). In this step, the polypeptides of this invention may serve as the immunogen without modification. Alternatively, particularly for relatively short polypeptides, a superior immune response may be elicited if the polypeptide is joined to a carrier protein, such as bovine serum albumin or keyhole limpet hemocyanin. The immunogen is injected into the animal host, preferably according to a predetermined schedule incorporating one or more booster immunizations, and the animals are bled periodically. Polyclonal antibodies specific for the polypeptide may then be purified from such antisera by, for example, affinity chromatography using the polypeptide coupled to a suitable solid support.

Monoclonal antibodies specific for an antigenic polypeptide of interest may be prepared, for example, using the technique of Kohler and Milstein, *Eur. J. Immunol.* 6:511-519, 1976, and improvements thereto. Briefly, these methods involve the preparation of immortal cell lines capable of producing antibodies having the desired specificity (*i.e.*, reactivity with the polypeptide of interest). Such cell lines may be produced, for example, from spleen cells obtained from an animal immunized as described above. The spleen cells are then immortalized by, for example, fusion with a myeloma cell fusion partner, preferably one that is syngeneic with the immunized animal. A variety of fusion techniques may be employed. For example, the spleen cells and myeloma cells may be combined with a nonionic detergent for a few minutes and then plated at low density on a selective medium that supports the growth of hybrid cells, but not myeloma cells. A preferred selection technique uses HAT (hypoxanthine, aminopterin, thymidine) selection. After a sufficient time, usually about 1 to 2 weeks, colonies of hybrids are observed. Single colonies are selected and their culture supernatants tested for binding activity against the polypeptide. Hybridomas having high reactivity and specificity are preferred.

Monoclonal antibodies may be isolated from the supernatants of growing hybridoma colonies. In addition, various techniques may be employed to enhance the yield, such as injection of the hybridoma cell line into the peritoneal cavity of a suitable vertebrate host, such as a mouse. Monoclonal antibodies may then be harvested from the ascites fluid or the blood. Contaminants may be removed from the antibodies by conventional techniques, such as chromatography, gel filtration, precipitation, and extraction. The polypeptides of this invention may be used in the purification process in, for example, an affinity chromatography step.

A number of therapeutically useful molecules are known in the art which comprise antigen-binding sites that are capable of exhibiting immunological binding properties of an antibody molecule. The proteolytic enzyme papain preferentially cleaves IgG molecules to yield several fragments, two of which (the "F(ab)" fragments) each comprise a covalent heterodimer that includes an intact antigen-binding site. The enzyme pepsin is able to cleave IgG molecules to provide several fragments, including the "F(ab)₂" fragment which comprises both antigen-binding sites. An "Fv" fragment can be produced by preferential proteolytic cleavage of an IgM, and on rare occasions IgG or IgA immunoglobulin molecule. Fv fragments are, however, more commonly derived using recombinant techniques known in the art. The Fv fragment includes a non-covalent V_H::V_L heterodimer including an antigen-binding site which retains much of the antigen recognition and binding capabilities of the native antibody molecule. Inbar et al. (1972) Proc. Nat. Acad. Sci. USA 69:2659-2662; Hochman et al. (1976) Biochem 15:2706-2710; and Ehrlich et al. (1980) Biochem 19:4091-4096.

A single chain Fv ("sFv") polypeptide is a covalently linked V_H::V_L heterodimer which is expressed from a gene fusion including V_H- and V_L-encoding genes linked by a peptide-encoding linker. Huston et al. (1988) Proc. Nat. Acad. Sci. USA 85(16):5879-5883. A number of methods have been described to discern chemical structures for converting the naturally aggregated--but chemically separated--light and heavy polypeptide chains from an antibody V region into an sFv molecule which will fold into a three dimensional structure substantially similar to the structure of an antigen-binding site. See, e.g., U.S. Pat. Nos. 5,091,513 and 5,132,405, to Huston et al.; and U.S. Pat. No. 4,946,778, to Ladner et al.

Each of the above-described molecules includes a heavy chain and a light chain CDR set, respectively interposed between a heavy chain and a light chain FR set which provide support to the CDRs and define the spatial relationship of the CDRs relative to each other. As used herein, the term "CDR set" refers to the three hypervariable regions of a heavy or light chain V region. Proceeding from the N-terminus of a heavy or light chain, these regions are denoted as "CDR1," "CDR2," and "CDR3" respectively. An antigen-binding site, therefore, includes six CDRs, comprising the CDR set from each of a heavy and a light chain V region. A polypeptide comprising a single CDR, (e.g., a CDR1, CDR2 or CDR3) is referred to herein as a "molecular recognition unit." Crystallographic analysis of a number of antigen-antibody complexes has demonstrated that the amino acid residues of CDRs form extensive contact with bound antigen, wherein the most extensive antigen contact is with the heavy chain CDR3. Thus, the molecular recognition units are primarily responsible for the specificity of an antigen-binding site.

As used herein, the term "FR set" refers to the four flanking amino acid sequences which frame the CDRs of a CDR set of a heavy or light chain V region. Some FR residues may contact bound antigen; however, FRs are primarily responsible for folding the V region into the antigen-binding site, particularly the FR residues directly adjacent to the CDRs. Within FRs, certain amino residues and certain structural features are very highly conserved. In this regard, all V region sequences contain an internal disulfide loop of around 90 amino acid residues. When the V regions fold into a binding-site, the CDRs are displayed as projecting loop motifs which form an antigen-binding surface. It is generally recognized that there are conserved structural regions of FRs which influence the folded shape of the CDR loops into certain "canonical" structures--regardless of the precise CDR amino acid sequence. Further, certain FR residues are known to participate in non-covalent interdomain contacts which stabilize the interaction of the antibody heavy and light chains.

A number of "humanized" antibody molecules comprising an antigen-binding site derived from a non-human immunoglobulin have been described, including chimeric antibodies having rodent V regions and their associated CDRs fused to human constant domains (Winter et al. (1991) Nature 349:293-299; Lobuglio et al. (1989)

Proc. Nat. Acad. Sci. USA 86:4220-4224; Shaw et al. (1987) J Immunol. 138:4534-4538; and Brown et al. (1987) Cancer Res. 47:3577-3583), rodent CDRs grafted into a human supporting FR prior to fusion with an appropriate human antibody constant domain (Riechmann et al. (1988) Nature 332:323-327; Verhoeyen et al. (1988) Science 5 239:1534-1536; and Jones et al. (1986) Nature 321:522-525), and rodent CDRs supported by recombinantly veneered rodent FRs (European Patent Publication No. 519,596, published Dec. 23, 1992). These "humanized" molecules are designed to minimize unwanted immunological response toward rodent antihuman antibody molecules which limits the duration and effectiveness of therapeutic applications of 10 those moieties in human recipients.

As used herein, the terms "veneered FRs" and "recombinantly veneered FRs" refer to the selective replacement of FR residues from, *e.g.*, a rodent heavy or light chain V region, with human FR residues in order to provide a xenogeneic molecule comprising an antigen-binding site which retains substantially all of the native FR 15 polypeptide folding structure. Veneering techniques are based on the understanding that the ligand binding characteristics of an antigen-binding site are determined primarily by the structure and relative disposition of the heavy and light chain CDR sets within the antigen-binding surface. Davies et al. (1990) Ann. Rev. Biochem. 59:439-473. Thus, antigen binding specificity can be preserved in a humanized antibody only wherein the 20 CDR structures, their interaction with each other, and their interaction with the rest of the V region domains are carefully maintained. By using veneering techniques, exterior (*e.g.*, solvent-accessible) FR residues which are readily encountered by the immune system are selectively replaced with human residues to provide a hybrid molecule that comprises either a weakly immunogenic, or substantially non-immunogenic veneered 25 surface.

The process of veneering makes use of the available sequence data for human antibody variable domains compiled by Kabat et al., in Sequences of Proteins of Immunological Interest, 4th ed., (U.S. Dept. of Health and Human Services, U.S. Government Printing Office, 1987), updates to the Kabat database, and other accessible 30 U.S. and foreign databases (both nucleic acid and protein). Solvent accessibilities of V region amino acids can be deduced from the known three-dimensional structure for

human and murine antibody fragments. There are two general steps in veneering a murine antigen-binding site. Initially, the FRs of the variable domains of an antibody molecule of interest are compared with corresponding FR sequences of human variable domains obtained from the above-identified sources. The most homologous human V regions are then compared residue by residue to corresponding murine amino acids. The residues in the murine FR which differ from the human counterpart are replaced by the residues present in the human moiety using recombinant techniques well known in the art. Residue switching is only carried out with moieties which are at least partially exposed (solvent accessible), and care is exercised in the replacement of amino acid residues which may have a significant effect on the tertiary structure of V region domains, such as proline, glycine and charged amino acids.

In this manner, the resultant "veneered" murine antigen-binding sites are thus designed to retain the murine CDR residues, the residues substantially adjacent to the CDRs, the residues identified as buried or mostly buried (solvent inaccessible), the residues believed to participate in non-covalent (e.g., electrostatic and hydrophobic) contacts between heavy and light chain domains, and the residues from conserved structural regions of the FRs which are believed to influence the "canonical" tertiary structures of the CDR loops. These design criteria are then used to prepare recombinant nucleotide sequences which combine the CDRs of both the heavy and light chain of a murine antigen-binding site into human-appearing FRs that can be used to transfect mammalian cells for the expression of recombinant human antibodies which exhibit the antigen specificity of the murine antibody molecule.

In another embodiment of the invention, monoclonal antibodies of the present invention may be coupled to one or more therapeutic agents. Suitable agents in this regard include radionuclides, differentiation inducers, drugs, toxins, and derivatives thereof. Preferred radionuclides include ^{90}Y , ^{123}I , ^{125}I , ^{131}I , ^{186}Re , ^{188}Re , ^{211}At , and ^{212}Bi . Preferred drugs include methotrexate, and pyrimidine and purine analogs. Preferred differentiation inducers include phorbol esters and butyric acid. Preferred toxins include ricin, abrin, diphtheria toxin, cholera toxin, gelonin, *Pseudomonas* exotoxin, Shigella toxin, and pokeweed antiviral protein.

A therapeutic agent may be coupled (*e.g.*, covalently bonded) to a suitable monoclonal antibody either directly or indirectly (*e.g.*, via a linker group). A direct reaction between an agent and an antibody is possible when each possesses a substituent capable of reacting with the other. For example, a nucleophilic group, such as an amino or sulfhydryl group, on one may be capable of reacting with a carbonyl-containing group, such as an anhydride or an acid halide, or with an alkyl group containing a good leaving group (*e.g.*, a halide) on the other.

Alternatively, it may be desirable to couple a therapeutic agent and an antibody via a linker group. A linker group can function as a spacer to distance an antibody from an agent in order to avoid interference with binding capabilities. A linker group can also serve to increase the chemical reactivity of a substituent on an agent or an antibody, and thus increase the coupling efficiency. An increase in chemical reactivity may also facilitate the use of agents, or functional groups on agents, which otherwise would not be possible.

It will be evident to those skilled in the art that a variety of bifunctional or polyfunctional reagents, both homo- and hetero-functional (such as those described in the catalog of the Pierce Chemical Co., Rockford, IL), may be employed as the linker group. Coupling may be effected, for example, through amino groups, carboxyl groups, sulfhydryl groups or oxidized carbohydrate residues. There are numerous references describing such methodology, *e.g.*, U.S. Patent No. 4,671,958, to Rodwell et al.

Where a therapeutic agent is more potent when free from the antibody portion of the immunoconjugates of the present invention, it may be desirable to use a linker group which is cleavable during or upon internalization into a cell. A number of different cleavable linker groups have been described. The mechanisms for the intracellular release of an agent from these linker groups include cleavage by reduction of a disulfide bond (*e.g.*, U.S. Patent No. 4,489,710, to Spitler), by irradiation of a photolabile bond (*e.g.*, U.S. Patent No. 4,625,014, to Senter et al.), by hydrolysis of derivatized amino acid side chains (*e.g.*, U.S. Patent No. 4,638,045, to Kohn et al.), by serum complement-mediated hydrolysis (*e.g.*, U.S. Patent No. 4,671,958, to Rodwell et al.), and acid-catalyzed hydrolysis (*e.g.*, U.S. Patent No. 4,569,789, to Blattler et al.).

It may be desirable to couple more than one agent to an antibody. In one embodiment, multiple molecules of an agent are coupled to one antibody molecule. In another embodiment, more than one type of agent may be coupled to one antibody. Regardless of the particular embodiment, immunoconjugates with more than one agent
5 may be prepared in a variety of ways. For example, more than one agent may be coupled directly to an antibody molecule, or linkers that provide multiple sites for attachment can be used. Alternatively, a carrier can be used.

A carrier may bear the agents in a variety of ways, including covalent bonding either directly or via a linker group. Suitable carriers include proteins such as
10 albumins (*e.g.*, U.S. Patent No. 4,507,234, to Kato et al.), peptides and polysaccharides such as aminodextran (*e.g.*, U.S. Patent No. 4,699,784, to Shih et al.). A carrier may also bear an agent by noncovalent bonding or by encapsulation, such as within a liposome vesicle (*e.g.*, U.S. Patent Nos. 4,429,008 and 4,873,088). Carriers specific for radionuclide agents include radiohalogenated small molecules and chelating
15 compounds. For example, U.S. Patent No. 4,735,792 discloses representative radiohalogenated small molecules and their synthesis. A radionuclide chelate may be formed from chelating compounds that include those containing nitrogen and sulfur atoms as the donor atoms for binding the metal, or metal oxide, radionuclide. For example, U.S. Patent No. 4,673,562, to Davison et al. discloses representative chelating
20 compounds and their synthesis.

T Cell Compositions

The present invention, in another aspect, provides T cells specific for a tumor polypeptide disclosed herein, or for a variant or derivative thereof. Such cells may generally be prepared *in vitro* or *ex vivo*, using standard procedures. For example,
25 T cells may be isolated from bone marrow, peripheral blood, or a fraction of bone marrow or peripheral blood of a patient, using a commercially available cell separation system, such as the Isolex™ System, available from Nexell Therapeutics, Inc. (Irvine, CA; see also U.S. Patent No. 5,240,856; U.S. Patent No. 5,215,926; WO 89/06280; WO 91/16116 and WO 92/07243). Alternatively, T cells may be derived from related or
30 unrelated humans, non-human mammals, cell lines or cultures.

T cells may be stimulated with a polypeptide, polynucleotide encoding a polypeptide and/or an antigen presenting cell (APC) that expresses such a polypeptide. Such stimulation is performed under conditions and for a time sufficient to permit the generation of T cells that are specific for the polypeptide of interest. Preferably, a tumor
5 polypeptide or polynucleotide of the invention is present within a delivery vehicle, such as a microsphere, to facilitate the generation of specific T cells.

T cells are considered to be specific for a polypeptide of the present invention if the T cells specifically proliferate, secrete cytokines or kill target cells coated with the polypeptide or expressing a gene encoding the polypeptide. T cell
10 specificity may be evaluated using any of a variety of standard techniques. For example, within a chromium release assay or proliferation assay, a stimulation index of more than two fold increase in lysis and/or proliferation, compared to negative controls, indicates T cell specificity. Such assays may be performed, for example, as described in Chen et al., *Cancer Res.* 54:1065-1070, 1994. Alternatively, detection of the
15 proliferation of T cells may be accomplished by a variety of known techniques. For example, T cell proliferation can be detected by measuring an increased rate of DNA synthesis (e.g., by pulse-labeling cultures of T cells with tritiated thymidine and measuring the amount of tritiated thymidine incorporated into DNA). Contact with a tumor polypeptide (100 ng/ml - 100 µg/ml, preferably 200 ng/ml - 25 µg/ml) for 3 - 7
20 days will typically result in at least a two fold increase in proliferation of the T cells. Contact as described above for 2-3 hours should result in activation of the T cells, as measured using standard cytokine assays in which a two fold increase in the level of cytokine release (e.g., TNF or IFN-γ) is indicative of T cell activation (see Coligan et al., *Current Protocols in Immunology*, vol. 1, Wiley Interscience (Greene 1998)). T
25 cells that have been activated in response to a tumor polypeptide, polynucleotide or polypeptide-expressing APC may be CD4⁺ and/or CD8⁺. Tumor polypeptide-specific T cells may be expanded using standard techniques. Within preferred embodiments, the T cells are derived from a patient, a related donor or an unrelated donor, and are administered to the patient following stimulation and expansion.

30 For therapeutic purposes, CD4⁺ or CD8⁺ T cells that proliferate in response to a tumor polypeptide, polynucleotide or APC can be expanded in number

either *in vitro* or *in vivo*. Proliferation of such T cells *in vitro* may be accomplished in a variety of ways. For example, the T cells can be re-exposed to a tumor polypeptide, or a short peptide corresponding to an immunogenic portion of such a polypeptide, with or without the addition of T cell growth factors, such as interleukin-2, and/or stimulator
5 cells that synthesize a tumor polypeptide. Alternatively, one or more T cells that proliferate in the presence of the tumor polypeptide can be expanded in number by cloning. Methods for cloning cells are well known in the art, and include limiting dilution.

T Cell Receptor Compositions

10 The T cell receptor (TCR) consists of 2 different, highly variable polypeptide chains, termed the T-cell receptor α and β chains, that are linked by a disulfide bond (Janeway, Travers, Walport. *Immunobiology*. Fourth Ed., 148-159. Elsevier Science Ltd/Garland Publishing. 1999). The α/β heterodimer complexes with the invariant CD3 chains at the cell membrane. This complex recognizes specific
15 antigenic peptides bound to MHC molecules. The enormous diversity of TCR specificities is generated much like immunoglobulin diversity, through somatic gene rearrangement. The β chain genes contain over 50 variable (V), 2 diversity (D), over 10 joining (J) segments, and 2 constant region segments (C). The α chain genes contain over 70 V segments, and over 60 J segments but no D segments, as well as one C
20 segment. During T cell development in the thymus, the D to J gene rearrangement of the β chain occurs, followed by the V gene segment rearrangement to the DJ. This functional VDJ $_{\beta}$ exon is transcribed and spliced to join to a C $_{\beta}$. For the α chain, a V $_{\alpha}$ gene segment rearranges to a J $_{\alpha}$ gene segment to create the functional exon that is then transcribed and spliced to the C $_{\alpha}$. Diversity is further increased during the
25 recombination process by the random addition of P and N-nucleotides between the V, D, and J segments of the β chain and between the V and J segments in the α chain (Janeway, Travers, Walport. *Immunobiology*. Fourth Ed., 98 and 150. Elsevier Science Ltd/Garland Publishing. 1999).

The present invention, in another aspect, provides TCRs specific for a polypeptide disclosed herein, or for a variant or derivative thereof. In accordance with the present invention, polynucleotide and amino acid sequences are provided for the V-J or V-D-J junctional regions or parts thereof for the alpha and beta chains of the T-cell receptor which recognize tumor polypeptides described herein. In general, this aspect of the invention relates to T-cell receptors which recognize or bind tumor polypeptides presented in the context of MHC. In a preferred embodiment the tumor antigens recognized by the T-cell receptors comprise a polypeptide of the present invention. For example, cDNA encoding a TCR specific for a tumor peptide can be isolated from T cells specific for a tumor polypeptide using standard molecular biological and recombinant DNA techniques.

This invention further includes the T-cell receptors or analogs thereof having substantially the same function or activity as the T-cell receptors of this invention which recognize or bind tumor polypeptides. Such receptors include, but are not limited to, a fragment of the receptor, or a substitution, addition or deletion mutant of a T-cell receptor provided herein. This invention also encompasses polypeptides or peptides that are substantially homologous to the T-cell receptors provided herein or that retain substantially the same activity. The term "analog" includes any protein or polypeptide having an amino acid residue sequence substantially identical to the T-cell receptors provided herein in which one or more residues, preferably no more than 5 residues, more preferably no more than 25 residues have been conservatively substituted with a functionally similar residue and which displays the functional aspects of the T-cell receptor as described herein.

The present invention further provides for suitable mammalian host cells, for example, non-specific T cells, that are transfected with a polynucleotide encoding TCRs specific for a polypeptide described herein, thereby rendering the host cell specific for the polypeptide. The α and β chains of the TCR may be contained on separate expression vectors or alternatively, on a single expression vector that also contains an internal ribosome entry site (IRES) for cap-independent translation of the gene downstream of the IRES. Said host cells expressing TCRs specific for the

polypeptide may be used, for example, for adoptive immunotherapy of ovarian cancer as discussed further below.

In further aspects of the present invention, cloned TCRs specific for a polypeptide recited herein may be used in a kit for the diagnosis of ovarian cancer. For example, the nucleic acid sequence or portions thereof, of tumor-specific TCRs can be used as probes or primers for the detection of expression of the rearranged genes encoding the specific TCR in a biological sample. Therefore, the present invention further provides for an assay for detecting messenger RNA or DNA encoding the TCR specific for a polypeptide.

10 Pharmaceutical Compositions

In additional embodiments, the present invention concerns formulation of one or more of the polynucleotide, polypeptide, T-cell, TCR, and/or antibody compositions disclosed herein in pharmaceutically-acceptable carriers for administration to a cell or an animal, either alone, or in combination with one or more other modalities of therapy.

It will be understood that, if desired, a composition as disclosed herein may be administered in combination with other agents as well, such as, *e.g.*, other proteins or polypeptides or various pharmaceutically-active agents. In fact, there is virtually no limit to other components that may also be included, given that the additional agents do not cause a significant adverse effect upon contact with the target cells or host tissues. The compositions may thus be delivered along with various other agents as required in the particular instance. Such compositions may be purified from host cells or other biological sources, or alternatively may be chemically synthesized as described herein. Likewise, such compositions may further comprise substituted or derivatized RNA or DNA compositions.

Therefore, in another aspect of the present invention, pharmaceutical compositions are provided comprising one or more of the polynucleotide, polypeptide, antibody, TCR, and/or T-cell compositions described herein in combination with a physiologically acceptable carrier. In certain preferred embodiments, the

pharmaceutical compositions of the invention comprise immunogenic polynucleotide and/or polypeptide compositions of the invention for use in prophylactic and therapeutic vaccine applications. Vaccine preparation is generally described in, for example, M.F. Powell and M.J. Newman, eds., "Vaccine Design (the subunit and adjuvant approach)," 5 Plenum Press (NY, 1995). Generally, such compositions will comprise one or more polynucleotide and/or polypeptide compositions of the present invention in combination with one or more immunostimulants.

It will be apparent that any of the pharmaceutical compositions described herein can contain pharmaceutically acceptable salts of the polynucleotides and 10 polypeptides of the invention. Such salts can be prepared, for example, from pharmaceutically acceptable non-toxic bases, including organic bases (e.g., salts of primary, secondary and tertiary amines and basic amino acids) and inorganic bases (e.g., sodium, potassium, lithium, ammonium, calcium and magnesium salts).

In another embodiment, illustrative immunogenic compositions, e.g., 15 vaccine compositions, of the present invention comprise DNA encoding one or more of the polypeptides as described above, such that the polypeptide is generated *in situ*. As noted above, the polynucleotide may be administered within any of a variety of delivery systems known to those of ordinary skill in the art. Indeed, numerous gene delivery techniques are well known in the art, such as those described by Rolland, *Crit. Rev.* 20 *Therap. Drug Carrier Systems* 15:143-198, 1998, and references cited therein. Appropriate polynucleotide expression systems will, of course, contain the necessary regulatory DNA regulatory sequences for expression in a patient (such as a suitable promoter and terminating signal). Alternatively, bacterial delivery systems may involve the administration of a bacterium (such as *Bacillus-Calmette-Guerrin*) that expresses an 25 immunogenic portion of the polypeptide on its cell surface or secretes such an epitope.

Therefore, in certain embodiments, polynucleotides encoding immunogenic polypeptides described herein are introduced into suitable mammalian host cells for expression using any of a number of known viral-based systems. In one illustrative embodiment, retroviruses provide a convenient and effective platform for 30 gene delivery systems. A selected nucleotide sequence encoding a polypeptide of the present invention can be inserted into a vector and packaged in retroviral particles using

techniques known in the art. The recombinant virus can then be isolated and delivered to a subject. A number of illustrative retroviral systems have been described (*e.g.*, U.S. Pat. No. 5,219,740; Miller and Rosman (1989) *BioTechniques* 7:980-990; Miller, A. D. (1990) *Human Gene Therapy* 1:5-14; Scarpa et al. (1991) *Virology* 180:849-852; Burns et al. (1993) *Proc. Natl. Acad. Sci. USA* 90:8033-8037; and Boris-Lawrie and Temin (1993) *Cur. Opin. Genet. Develop.* 3:102-109.

In addition, a number of illustrative adenovirus-based systems have also been described. Unlike retroviruses which integrate into the host genome, adenoviruses persist extrachromosomally thus minimizing the risks associated with insertional mutagenesis (Haj-Ahmad and Graham (1986) *J. Virol.* 57:267-274; Bett et al. (1993) *J. Virol.* 67:5911-5921; Mittereder et al. (1994) *Human Gene Therapy* 5:717-729; Seth et al. (1994) *J. Virol.* 68:933-940; Barr et al. (1994) *Gene Therapy* 1:51-58; Berkner, K. L. (1988) *BioTechniques* 6:616-629; and Rich et al. (1993) *Human Gene Therapy* 4:461-476).

Various adeno-associated virus (AAV) vector systems have also been developed for polynucleotide delivery. AAV vectors can be readily constructed using techniques well known in the art. See, *e.g.*, U.S. Pat. Nos. 5,173,414 and 5,139,941; International Publication Nos. WO 92/01070 and WO 93/03769; Lebkowski et al. (1988) *Molec. Cell. Biol.* 8:3988-3996; Vincent et al. (1990) *Vaccines* 90 (Cold Spring Harbor Laboratory Press); Carter, B. J. (1992) *Current Opinion in Biotechnology* 3:533-539; Muzyczka, N. (1992) *Current Topics in Microbiol. and Immunol.* 158:97-129; Kotin, R. M. (1994) *Human Gene Therapy* 5:793-801; Shelling and Smith (1994) *Gene Therapy* 1:165-169; and Zhou et al. (1994) *J. Exp. Med.* 179:1867-1875.

Additional viral vectors useful for delivering the polynucleotides encoding polypeptides of the present invention by gene transfer include those derived from the pox family of viruses, such as vaccinia virus and avian poxvirus. By way of example, vaccinia virus recombinants expressing the novel molecules can be constructed as follows. The DNA encoding a polypeptide is first inserted into an appropriate vector so that it is adjacent to a vaccinia promoter and flanking vaccinia DNA sequences, such as the sequence encoding thymidine kinase (TK). This vector is then used to transfect cells which are simultaneously infected with vaccinia.

Homologous recombination serves to insert the vaccinia promoter plus the gene encoding the polypeptide of interest into the viral genome. The resulting TK^{sup}(-) recombinant can be selected by culturing the cells in the presence of 5-bromodeoxyuridine and picking viral plaques resistant thereto.

5 A vaccinia-based infection/transfection system can be conveniently used to provide for inducible, transient expression or coexpression of one or more polypeptides described herein in host cells of an organism. In this particular system, cells are first infected in vitro with a vaccinia virus recombinant that encodes the bacteriophage T7 RNA polymerase. This polymerase displays exquisite specificity in
10 that it only transcribes templates bearing T7 promoters. Following infection, cells are transfected with the polynucleotide or polynucleotides of interest, driven by a T7 promoter. The polymerase expressed in the cytoplasm from the vaccinia virus recombinant transcribes the transfected DNA into RNA which is then translated into polypeptide by the host translational machinery. The method provides for high level,
15 transient, cytoplasmic production of large quantities of RNA and its translation products. See, *e.g.*, Elroy-Stein and Moss, Proc. Natl. Acad. Sci. USA (1990) 87:6743-6747; Fuerst et al. Proc. Natl. Acad. Sci. USA (1986) 83:8122-8126.

 Alternatively, avipoxviruses, such as the fowlpox and canarypox viruses, can also be used to deliver the coding sequences of interest. Recombinant avipox
20 viruses, expressing immunogens from mammalian pathogens, are known to confer protective immunity when administered to non-avian species. The use of an Avipox vector is particularly desirable in human and other mammalian species since members of the Avipox genus can only productively replicate in susceptible avian species and therefore are not infective in mammalian cells. Methods for producing recombinant
25 Avipoxviruses are known in the art and employ genetic recombination, as described above with respect to the production of vaccinia viruses. See, *e.g.*, WO 91/12882; WO 89/03429; and WO 92/03545.

 Any of a number of alphavirus vectors can also be used for delivery of polynucleotide compositions of the present invention, such as those vectors described in
30 U.S. Patent Nos. 5,843,723; 6,015,686; 6,008,035 and 6,015,694. Certain vectors based

on Venezuelan Equine Encephalitis (VEE) can also be used, illustrative examples of which can be found in U.S. Patent Nos. 5,505,947 and 5,643,576.

Moreover, molecular conjugate vectors, such as the adenovirus chimeric vectors described in Michael et al. *J. Biol. Chem.* (1993) 268:6866-6869 and Wagner et al. *Proc. Natl. Acad. Sci. USA* (1992) 89:6099-6103, can also be used for gene delivery under the invention.

Additional illustrative information on these and other known viral-based delivery systems can be found, for example, in Fisher-Hoch et al., *Proc. Natl. Acad. Sci. USA* 86:317-321, 1989; Flexner et al., *Ann. N.Y. Acad. Sci.* 569:86-103, 1989; Flexner et al., *Vaccine* 8:17-21, 1990; U.S. Patent Nos. 4,603,112, 4,769,330, and 5,017,487; WO 89/01973; U.S. Patent No. 4,777,127; GB 2,200,651; EP 0,345,242; WO 91/02805; Berkner, *Biotechniques* 6:616-627, 1988; Rosenfeld et al., *Science* 252:431-434, 1991; Kolls et al., *Proc. Natl. Acad. Sci. USA* 91:215-219, 1994; Kass-Eisler et al., *Proc. Natl. Acad. Sci. USA* 90:11498-11502, 1993; Guzman et al., *Circulation* 88:2838-2848, 1993; and Guzman et al., *Cir. Res.* 73:1202-1207, 1993.

In certain embodiments, a polynucleotide may be integrated into the genome of a target cell. This integration may be in the specific location and orientation via homologous recombination (gene replacement) or it may be integrated in a random, non-specific location (gene augmentation). In yet further embodiments, the polynucleotide may be stably maintained in the cell as a separate, episomal segment of DNA. Such polynucleotide segments or "episomes" encode sequences sufficient to permit maintenance and replication independent of or in synchronization with the host cell cycle. The manner in which the expression construct is delivered to a cell and where in the cell the polynucleotide remains is dependent on the type of expression construct employed.

In another embodiment of the invention, a polynucleotide is administered/delivered as "naked" DNA, for example as described in Ulmer et al., *Science* 259:1745-1749, 1993 and reviewed by Cohen, *Science* 259:1691-1692, 1993. The uptake of naked DNA may be increased by coating the DNA onto biodegradable beads, which are efficiently transported into the cells.

In still another embodiment, a composition of the present invention can be delivered via a particle bombardment approach, many of which have been described. In one illustrative example, gas-driven particle acceleration can be achieved with devices such as those manufactured by Powderject Pharmaceuticals PLC (Oxford, UK) and Powderject Vaccines Inc. (Madison, WI), some examples of which are described in U.S. Patent Nos. 5,846,796; 6,010,478; 5,865,796; 5,584,807; and EP Patent No. 0500 799. This approach offers a needle-free delivery approach wherein a dry powder formulation of microscopic particles, such as polynucleotide or polypeptide particles, are accelerated to high speed within a helium gas jet generated by a hand held device, propelling the particles into a target tissue of interest.

In a related embodiment, other devices and methods that may be useful for gas-driven needle-less injection of compositions of the present invention include those provided by Bioject, Inc. (Portland, OR), some examples of which are described in U.S. Patent Nos. 4,790,824; 5,064,413; 5,312,335; 5,383,851; 5,399,163; 5,520,639 and 5,993,412.

According to another embodiment, the pharmaceutical compositions described herein will comprise one or more immunostimulants in addition to the immunogenic polynucleotide, polypeptide, antibody, T-cell, TCR, and/or APC compositions of this invention. An immunostimulant refers to essentially any substance that enhances or potentiates an immune response (antibody and/or cell-mediated) to an exogenous antigen. One preferred type of immunostimulant comprises an adjuvant. Many adjuvants contain a substance designed to protect the antigen from rapid catabolism, such as aluminum hydroxide or mineral oil, and a stimulator of immune responses, such as lipid A, *Bordetella pertussis* or *Mycobacterium tuberculosis* derived proteins. Certain adjuvants are commercially available as, for example, Freund's Incomplete Adjuvant and Complete Adjuvant (Difco Laboratories, Detroit, MI); Merck Adjuvant 65 (Merck and Company, Inc., Rahway, NJ); AS-2 (SmithKline Beecham, Philadelphia, PA); aluminum salts such as aluminum hydroxide gel (alum) or aluminum phosphate; salts of calcium, iron or zinc; an insoluble suspension of acylated tyrosine; acylated sugars; cationically or anionically derivatized polysaccharides; polyphosphazenes; biodegradable microspheres; monophosphoryl lipid A and quil A.

Cytokines, such as GM-CSF, interleukin-2, -7, -12, and other like growth factors, may also be used as adjuvants.

Within certain embodiments of the invention, the adjuvant composition is preferably one that induces an immune response predominantly of the Th1 type. High levels of Th1-type cytokines (*e.g.*, IFN- γ , TNF α , IL-2 and IL-12) tend to favor the induction of cell mediated immune responses to an administered antigen. In contrast, high levels of Th2-type cytokines (*e.g.*, IL-4, IL-5, IL-6 and IL-10) tend to favor the induction of humoral immune responses. Following application of a vaccine as provided herein, a patient will support an immune response that includes Th1- and Th2-type responses. Within a preferred embodiment, in which a response is predominantly Th1-type, the level of Th1-type cytokines will increase to a greater extent than the level of Th2-type cytokines. The levels of these cytokines may be readily assessed using standard assays. For a review of the families of cytokines, see Mosmann and Coffman, *Ann. Rev. Immunol.* 7:145-173, 1989.

Certain preferred adjuvants for eliciting a predominantly Th1-type response include, for example, a combination of monophosphoryl lipid A, preferably 3-de-O-acylated monophosphoryl lipid A, together with an aluminum salt. MPL[®] adjuvants are available from Corixa Corporation (Seattle, WA; *see*, for example, US Patent Nos. 4,436,727; 4,877,611; 4,866,034 and 4,912,094). CpG-containing oligonucleotides (in which the CpG dinucleotide is unmethylated) also induce a predominantly Th1 response. Such oligonucleotides are well known and are described, for example, in WO 96/02555, WO 99/33488 and U.S. Patent Nos. 6,008,200 and 5,856,462. Immunostimulatory DNA sequences are also described, for example, by Sato et al., *Science* 273:352, 1996. Another preferred adjuvant comprises a saponin, such as Quil A, or derivatives thereof, including QS21 and QS7 (Aquila Biopharmaceuticals Inc., Framingham, MA); Escin; Digitonin; or *Gypsophila* or *Chenopodium quinoa* saponins. Other preferred formulations include more than one saponin in the adjuvant combinations of the present invention, for example combinations of at least two of the following group comprising QS21, QS7, Quil A, β -escin, or digitonin.

Alternatively the saponin formulations may be combined with vaccine vehicles composed of chitosan or other polycationic polymers, polylactide and polylactide-co-glycolide particles, poly-N-acetyl glucosamine-based polymer matrix, particles composed of polysaccharides or chemically modified polysaccharides, liposomes and lipid-based particles, particles composed of glycerol monoesters, etc. The saponins may also be formulated in the presence of cholesterol to form particulate structures such as liposomes or ISCOMs. Furthermore, the saponins may be formulated together with a polyoxyethylene ether or ester, in either a non-particulate solution or suspension, or in a particulate structure such as a paucilamellar liposome or ISCOM. The saponins may also be formulated with excipients such as Carbopol^R to increase viscosity, or may be formulated in a dry powder form with a powder excipient such as lactose.

In one preferred embodiment, the adjuvant system includes the combination of a monophosphoryl lipid A and a saponin derivative, such as the combination of QS21 and 3D-MPL[®] adjuvant, as described in WO 94/00153, or a less reactogenic composition where the QS21 is quenched with cholesterol, as described in WO 96/33739. Other preferred formulations comprise an oil-in-water emulsion and tocopherol. Another particularly preferred adjuvant formulation employing QS21, 3D-MPL[®] adjuvant and tocopherol in an oil-in-water emulsion is described in WO 95/17210.

Another enhanced adjuvant system involves the combination of a CpG-containing oligonucleotide and a saponin derivative particularly the combination of CpG and QS21 is disclosed in WO 00/09159. Preferably the formulation additionally comprises an oil in water emulsion and tocopherol.

Additional illustrative adjuvants for use in the pharmaceutical compositions of the invention include Montanide ISA 720 (Seppic, France), SAF (Chiron, California, United States), ISCOMS (CSL), MF-59 (Chiron), the SBAS series of adjuvants (e.g., SBAS-2 or SBAS-4, available from SmithKline Beecham, Rixensart, Belgium), Detox (Enhanzyn[®]) (Corixa, Hamilton, MT), RC-529 (Corixa, Hamilton, MT) and other aminoalkyl glucosaminide 4-phosphates (AGPs), such as those described in pending U.S. Patent Application Serial Nos. 08/853,826 and 09/074,720, the

disclosures of which are incorporated herein by reference in their entireties, and polyoxyethylene ether adjuvants such as those described in WO 99/52549A1.

Other preferred adjuvants include adjuvant molecules of the general formula

5 (I): $\text{HO}(\text{CH}_2\text{CH}_2\text{O})_n\text{-A-R}$,

wherein, n is 1-50, A is a bond or $-\text{C}(\text{O})-$, R is C_{1-50} alkyl or Phenyl C_{1-50} alkyl.

One embodiment of the present invention consists of a vaccine formulation comprising a polyoxyethylene ether of general formula (I), wherein n is between 1 and 50, preferably 4-24, most preferably 9; the R component is C_{1-50} ,
10 preferably $\text{C}_4\text{-C}_{20}$ alkyl and most preferably C_{12} alkyl, and A is a bond. The concentration of the polyoxyethylene ethers should be in the range 0.1-20%, preferably from 0.1-10%, and most preferably in the range 0.1-1%. Preferred polyoxyethylene ethers are selected from the following group: polyoxyethylene-9-lauryl ether, polyoxyethylene-9-stearyl ether, polyoxyethylene-8-stearyl ether, polyoxyethylene-4-
15 lauryl ether, polyoxyethylene-35-lauryl ether, and polyoxyethylene-23-lauryl ether. Polyoxyethylene ethers such as polyoxyethylene lauryl ether are described in the Merck index (12th edition: entry 7717). These adjuvant molecules are described in WO 99/52549.

The polyoxyethylene ether according to the general formula (I) above
20 may, if desired, be combined with another adjuvant. For example, a preferred adjuvant combination is preferably with CpG as described in the pending UK patent application GB 9820956.2.

According to another embodiment of this invention, an immunogenic composition described herein is delivered to a host via antigen presenting cells (APCs),
25 such as dendritic cells, macrophages, B cells, monocytes and other cells that may be engineered to be efficient APCs. Such cells may, but need not, be genetically modified to increase the capacity for presenting the antigen, to improve activation and/or maintenance of the T cell response, to have anti-tumor effects *per se* and/or to be immunologically compatible with the receiver (*i.e.*, matched HLA haplotype). APCs
30 may generally be isolated from any of a variety of biological fluids and organs,

including tumor and peritumoral tissues, and may be autologous, allogeneic, syngeneic or xenogeneic cells.

Certain preferred embodiments of the present invention use dendritic cells or progenitors thereof as antigen-presenting cells. Dendritic cells are highly potent
5 APCs (Banchereau and Steinman, *Nature* 392:245-251, 1998) and have been shown to be effective as a physiological adjuvant for eliciting prophylactic or therapeutic antitumor immunity (*see* Timmerman and Levy, *Ann. Rev. Med.* 50:507-529, 1999). In general, dendritic cells may be identified based on their typical shape (stellate *in situ*, with marked cytoplasmic processes (dendrites) visible *in vitro*), their ability to take up,
10 process and present antigens with high efficiency and their ability to activate naïve T cell responses. Dendritic cells may, of course, be engineered to express specific cell-surface receptors or ligands that are not commonly found on dendritic cells *in vivo* or *ex vivo*, and such modified dendritic cells are contemplated by the present invention. As an alternative to dendritic cells, secreted vesicles antigen-loaded dendritic cells (called
15 exosomes) may be used within a vaccine (*see* Zitvogel et al., *Nature Med.* 4:594-600, 1998).

Dendritic cells and progenitors may be obtained from peripheral blood, bone marrow, tumor-infiltrating cells, peritumoral tissues-infiltrating cells, lymph nodes, spleen, skin, umbilical cord blood or any other suitable tissue or fluid. For
20 example, dendritic cells may be differentiated *ex vivo* by adding a combination of cytokines such as GM-CSF, IL-4, IL-13 and/or TNF α to cultures of monocytes harvested from peripheral blood. Alternatively, CD34 positive cells harvested from peripheral blood, umbilical cord blood or bone marrow may be differentiated into dendritic cells by adding to the culture medium combinations of GM-CSF, IL-3, TNF α ,
25 CD40 ligand, LPS, flt3 ligand and/or other compound(s) that induce differentiation, maturation and proliferation of dendritic cells.

Dendritic cells are conveniently categorized as "immature" and "mature" cells, which allows a simple way to discriminate between two well characterized phenotypes. However, this nomenclature should not be construed to exclude all
30 possible intermediate stages of differentiation. Immature dendritic cells are

characterized as APC with a high capacity for antigen uptake and processing, which correlates with the high expression of Fcγ receptor and mannose receptor. The mature phenotype is typically characterized by a lower expression of these markers, but a high expression of cell surface molecules responsible for T cell activation such as class I and class II MHC, adhesion molecules (e.g., CD54 and CD11) and costimulatory molecules (e.g., CD40, CD80, CD86 and 4-1BB).

APCs may generally be transfected with a polynucleotide of the invention (or portion or other variant thereof) such that the encoded polypeptide, or an immunogenic portion thereof, is expressed on the cell surface. Such transfection may take place *ex vivo*, and a pharmaceutical composition comprising such transfected cells may then be used for therapeutic purposes, as described herein. Alternatively, a gene delivery vehicle that targets a dendritic or other antigen presenting cell may be administered to a patient, resulting in transfection that occurs *in vivo*. *In vivo* and *ex vivo* transfection of dendritic cells, for example, may generally be performed using any methods known in the art, such as those described in WO 97/24447, or the gene gun approach described by Mahvi et al., *Immunology and cell Biology* 75:456-460, 1997. Antigen loading of dendritic cells may be achieved by incubating dendritic cells or progenitor cells with the tumor polypeptide, DNA (naked or within a plasmid vector) or RNA; or with antigen-expressing recombinant bacterium or viruses (e.g., vaccinia, fowlpox, adenovirus or lentivirus vectors). Prior to loading, the polypeptide may be covalently conjugated to an immunological partner that provides T cell help (e.g., a carrier molecule). Alternatively, a dendritic cell may be pulsed with a non-conjugated immunological partner, separately or in the presence of the polypeptide.

While any suitable carrier known to those of ordinary skill in the art may be employed in the pharmaceutical compositions of this invention, the type of carrier will typically vary depending on the mode of administration. Compositions of the present invention may be formulated for any appropriate manner of administration, including for example, topical, oral, nasal, mucosal, intravenous, intracranial, intraperitoneal, subcutaneous and intramuscular administration.

Carriers for use within such pharmaceutical compositions are biocompatible, and may also be biodegradable. In certain embodiments, the

formulation preferably provides a relatively constant level of active component release. In other embodiments, however, a more rapid rate of release immediately upon administration may be desired. The formulation of such compositions is well within the level of ordinary skill in the art using known techniques. Illustrative carriers useful in this regard include microparticles of poly(lactide-co-glycolide), polyacrylate, latex, starch, cellulose, dextran and the like. Other illustrative delayed-release carriers include supramolecular biovectors, which comprise a non-liquid hydrophilic core (*e.g.*, a cross-linked polysaccharide or oligosaccharide) and, optionally, an external layer comprising an amphiphilic compound, such as a phospholipid (*see e.g.*, U.S. Patent No. 5,151,254 and PCT applications WO 94/20078, WO/94/23701 and WO 96/06638). The amount of active compound contained within a sustained release formulation depends upon the site of implantation, the rate and expected duration of release and the nature of the condition to be treated or prevented.

In another illustrative embodiment, biodegradable microspheres (*e.g.*, polylactate polyglycolate) are employed as carriers for the compositions of this invention. Suitable biodegradable microspheres are disclosed, for example, in U.S. Patent Nos. 4,897,268; 5,075,109; 5,928,647; 5,811,128; 5,820,883; 5,853,763; 5,814,344, 5,407,609 and 5,942,252. Modified hepatitis B core protein carrier systems, such as described in WO/99 40934, and references cited therein, will also be useful for many applications. Another illustrative carrier/delivery system employs a carrier comprising particulate-protein complexes, such as those described in U.S. Patent No. 5,928,647, which are capable of inducing a class I-restricted cytotoxic T lymphocyte responses in a host.

In another illustrative embodiment, calcium phosphate core particles are employed as carriers, vaccine adjuvants, or as controlled release matrices for the compositions of this invention. Exemplary calcium phosphate particles are disclosed, for example, in published patent application No. WO/0046147.

The pharmaceutical compositions of the invention will often further comprise one or more buffers (*e.g.*, neutral buffered saline or phosphate buffered saline), carbohydrates (*e.g.*, glucose, mannose, sucrose or dextrans), mannitol, proteins, polypeptides or amino acids such as glycine, antioxidants, bacteriostats, chelating

agents such as EDTA or glutathione, adjuvants (*e.g.*, aluminum hydroxide), solutes that render the formulation isotonic, hypotonic or weakly hypertonic with the blood of a recipient, suspending agents, thickening agents and/or preservatives. Alternatively, compositions of the present invention may be formulated as a lyophilizate.

5 The pharmaceutical compositions described herein may be presented in unit-dose or multi-dose containers, such as sealed ampoules or vials. Such containers are typically sealed in such a way to preserve the sterility and stability of the formulation until use. In general, formulations may be stored as suspensions, solutions or emulsions in oily or aqueous vehicles. Alternatively, a pharmaceutical composition
10 may be stored in a freeze-dried condition requiring only the addition of a sterile liquid carrier immediately prior to use.

 The development of suitable dosing and treatment regimens for using the particular compositions described herein in a variety of treatment regimens, including *e.g.*, oral, parenteral, intravenous, intranasal, and intramuscular administration and
15 formulation, is well known in the art, some of which are briefly discussed below for general purposes of illustration.

 In certain applications, the pharmaceutical compositions disclosed herein may be delivered *via* oral administration to an animal. As such, these compositions may be formulated with an inert diluent or with an assimilable edible carrier, or they
20 may be enclosed in hard- or soft-shell gelatin capsule, or they may be compressed into tablets, or they may be incorporated directly with the food of the diet.

 The active compounds may even be incorporated with excipients and used in the form of ingestible tablets, buccal tables, troches, capsules, elixirs, suspensions, syrups, wafers, and the like (see, for example, Mathiowitz *et al.*, Nature
25 1997 Mar 27;386(6623):410-4; Hwang *et al.*, Crit Rev Ther Drug Carrier Syst 1998;15(3):243-84; U. S. Patent 5,641,515; U. S. Patent 5,580,579 and U. S. Patent 5,792,451). Tablets, troches, pills, capsules and the like may also contain any of a variety of additional components, for example, a binder, such as gum tragacanth, acacia, cornstarch, or gelatin; excipients, such as dicalcium phosphate; a disintegrating agent,
30 such as corn starch, potato starch, alginic acid and the like; a lubricant, such as magnesium stearate; and a sweetening agent, such as sucrose, lactose or saccharin may

be added or a flavoring agent, such as peppermint, oil of wintergreen, or cherry flavoring. When the dosage unit form is a capsule, it may contain, in addition to materials of the above type, a liquid carrier. Various other materials may be present as coatings or to otherwise modify the physical form of the dosage unit. For instance, 5 tablets, pills, or capsules may be coated with shellac, sugar, or both. Of course, any material used in preparing any dosage unit form should be pharmaceutically pure and substantially non-toxic in the amounts employed. In addition, the active compounds may be incorporated into sustained-release preparation and formulations.

Typically, these formulations will contain at least about 0.1% of the 10 active compound or more, although the percentage of the active ingredient(s) may, of course, be varied and may conveniently be between about 1 or 2% and about 60% or 70% or more of the weight or volume of the total formulation. Naturally, the amount of active compound(s) in each therapeutically useful composition may be prepared in such a way that a suitable dosage will be obtained in any given unit dose of the compound. 15 Factors such as solubility, bioavailability, biological half-life, route of administration, product shelf life, as well as other pharmacological considerations will be contemplated by one skilled in the art of preparing such pharmaceutical formulations, and as such, a variety of dosages and treatment regimens may be desirable.

For oral administration the compositions of the present invention may 20 alternatively be incorporated with one or more excipients in the form of a mouthwash, dentifrice, buccal tablet, oral spray, or sublingual orally-administered formulation. Alternatively, the active ingredient may be incorporated into an oral solution such as one containing sodium borate, glycerin and potassium bicarbonate, or dispersed in a dentifrice, or added in a therapeutically-effective amount to a composition that may 25 include water, binders, abrasives, flavoring agents, foaming agents, and humectants. Alternatively the compositions may be fashioned into a tablet or solution form that may be placed under the tongue or otherwise dissolved in the mouth.

In certain circumstances it will be desirable to deliver the pharmaceutical compositions disclosed herein parenterally, intravenously, intramuscularly, or even 30 intraperitoneally. Such approaches are well known to the skilled artisan, some of which are further described, for example, in U. S. Patent 5,543,158; U. S. Patent 5,641,515

and U. S. Patent 5,399,363. In certain embodiments, solutions of the active compounds as free base or pharmacologically acceptable salts may be prepared in water suitably mixed with a surfactant, such as hydroxypropylcellulose. Dispersions may also be prepared in glycerol, liquid polyethylene glycols, and mixtures thereof and in oils.

- 5 Under ordinary conditions of storage and use, these preparations generally will contain a preservative to prevent the growth of microorganisms.

Illustrative pharmaceutical forms suitable for injectable use include sterile aqueous solutions or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersions (for example, see U. S. Patent
10 5,466,468). In all cases the form must be sterile and must be fluid to the extent that easy syringability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms, such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (*e.g.*, glycerol, propylene glycol, and
15 liquid polyethylene glycol, and the like), suitable mixtures thereof, and/or vegetable oils. Proper fluidity may be maintained, for example, by the use of a coating, such as lecithin, by the maintenance of the required particle size in the case of dispersion and/or by the use of surfactants. The prevention of the action of microorganisms can be facilitated by various antibacterial and antifungal agents, for example, parabens,
20 chlorobutanol, phenol, sorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars or sodium chloride. Prolonged absorption of the injectable compositions can be brought about by the use in the compositions of agents delaying absorption, for example, aluminum monostearate and gelatin.

- 25 In one embodiment, for parenteral administration in an aqueous solution, the solution should be suitably buffered if necessary and the liquid diluent first rendered isotonic with sufficient saline or glucose. These particular aqueous solutions are especially suitable for intravenous, intramuscular, subcutaneous and intraperitoneal administration. In this connection, a sterile aqueous medium that can be employed will
30 be known to those of skill in the art in light of the present disclosure. For example, one dosage may be dissolved in 1 ml of isotonic NaCl solution and either added to 1000 ml

of hypodermoclysis fluid or injected at the proposed site of infusion, (see for example, "Remington's Pharmaceutical Sciences" 15th Edition, pages 1035-1038 and 1570-1580). Some variation in dosage will necessarily occur depending on the condition of the subject being treated. Moreover, for human administration, preparations will of course preferably meet sterility, pyrogenicity, and the general safety and purity standards as required by FDA Office of Biologics standards.

In another embodiment of the invention, the compositions disclosed herein may be formulated in a neutral or salt form. Illustrative pharmaceutically-acceptable salts include the acid addition salts (formed with the free amino groups of the protein) and which are formed with inorganic acids such as, for example, hydrochloric or phosphoric acids, or such organic acids as acetic, oxalic, tartaric, mandelic, and the like. Salts formed with the free carboxyl groups can also be derived from inorganic bases such as, for example, sodium, potassium, ammonium, calcium, or ferric hydroxides, and such organic bases as isopropylamine, trimethylamine, histidine, procaine and the like. Upon formulation, solutions will be administered in a manner compatible with the dosage formulation and in such amount as is therapeutically effective.

The carriers can further comprise any and all solvents, dispersion media, vehicles, coatings, diluents, antibacterial and antifungal agents, isotonic and absorption delaying agents, buffers, carrier solutions, suspensions, colloids, and the like. The use of such media and agents for pharmaceutical active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active ingredient, its use in the therapeutic compositions is contemplated. Supplementary active ingredients can also be incorporated into the compositions. The phrase "pharmaceutically-acceptable" refers to molecular entities and compositions that do not produce an allergic or similar untoward reaction when administered to a human.

In certain embodiments, the pharmaceutical compositions may be delivered by intranasal sprays, inhalation, and/or other aerosol delivery vehicles. Methods for delivering genes, nucleic acids, and peptide compositions directly to the lungs via nasal aerosol sprays has been described, *e.g.*, in U. S. Patent 5,756,353 and U. S. Patent 5,804,212. Likewise, the delivery of drugs using intranasal microparticle

resins (Takenaga *et al.*, J Controlled Release 1998 Mar 2;52(1-2):81-7) and lysophosphatidyl-glycerol compounds (U. S. Patent 5,725,871) are also well-known in the pharmaceutical arts. Likewise, illustrative transmucosal drug delivery in the form of a polytetrafluoroethylene support matrix is described in U. S. Patent 5,780,045.

5 In certain embodiments, liposomes, nanocapsules, microparticles, lipid particles, vesicles, and the like, are used for the introduction of the compositions of the present invention into suitable host cells/organisms. In particular, the compositions of the present invention may be formulated for delivery either encapsulated in a lipid particle, a liposome, a vesicle, a nanosphere, or a nanoparticle or the like. Alternatively,
10 compositions of the present invention can be bound, either covalently or non-covalently, to the surface of such carrier vehicles.

 The formation and use of liposome and liposome-like preparations as potential drug carriers is generally known to those of skill in the art (see for example, Lasic, Trends Biotechnol 1998 Jul;16(7):307-21; Takakura, Nippon Rinsho 1998
15 Mar;56(3):691-5; Chandran *et al.*, Indian J Exp Biol. 1997 Aug;35(8):801-9; Margalit, Crit Rev Ther Drug Carrier Syst. 1995;12(2-3):233-61; U.S. Patent 5,567,434; U.S. Patent 5,552,157; U.S. Patent 5,565,213; U.S. Patent 5,738,868 and U.S. Patent 5,795,587, each specifically incorporated herein by reference in its entirety).

 Liposomes have been used successfully with a number of cell types that
20 are normally difficult to transfect by other procedures, including T cell suspensions, primary hepatocyte cultures and PC 12 cells (Renneisen *et al.*, J Biol Chem. 1990 Sep 25;265(27):16337-42; Muller *et al.*, DNA Cell Biol. 1990 Apr;9(3):221-9). In addition, liposomes are free of the DNA length constraints that are typical of viral-based delivery systems. Liposomes have been used effectively to introduce genes, various drugs,
25 radiotherapeutic agents, enzymes, viruses, transcription factors, allosteric effectors and the like, into a variety of cultured cell lines and animals. Furthermore, the use of liposomes does not appear to be associated with autoimmune responses or unacceptable toxicity after systemic delivery.

 In certain embodiments, liposomes are formed from phospholipids that
30 are dispersed in an aqueous medium and spontaneously form multilamellar concentric bilayer vesicles (also termed multilamellar vesicles (MLVs)).

Alternatively, in other embodiments, the invention provides for pharmaceutically-acceptable nanocapsule formulations of the compositions of the present invention. Nanocapsules can generally entrap compounds in a stable and reproducible way (see, for example, Quintanar-Guerrero *et al.*, Drug Dev Ind Pharm. 5 1998 Dec;24(12):1113-28). To avoid side effects due to intracellular polymeric overloading, such ultrafine particles (sized around 0.1 μm) may be designed using polymers able to be degraded *in vivo*. Such particles can be made as described, for example, by Couvreur *et al.*, Crit Rev Ther Drug Carrier Syst. 1988;5(1):1-20; zur Muhlen *et al.*, Eur J Pharm Biopharm. 1998 Mar;45(2):149-55; Zambaux *et al.* J 10 Controlled Release. 1998 Jan 2;50(1-3):31-40; and U. S. Patent 5,145,684.

Cancer Therapeutic Methods

Immunologic approaches to cancer therapy are based on the recognition that cancer cells can often evade the body's defenses against aberrant or foreign cells and molecules, and that these defenses might be therapeutically stimulated to regain the 15 lost ground, *e.g.* pgs. 623-648 in Klein, Immunology (Wiley-Interscience, New York, 1982). Numerous recent observations that various immune effectors can directly or indirectly inhibit growth of tumors has led to renewed interest in this approach to cancer therapy, *e.g.* Jager, et al., Oncology 2001;60(1):1-7; Renner, et al., Ann Hematol 2000 Dec;79(12):651-9.

20 Four-basic cell types whose function has been associated with antitumor cell immunity and the elimination of tumor cells from the body are: i) B-lymphocytes which secrete immunoglobulins into the blood plasma for identifying and labeling the nonself invader cells; ii) monocytes which secrete the complement proteins that are responsible for lysing and processing the immunoglobulin-coated target invader cells; 25 iii) natural killer lymphocytes having two mechanisms for the destruction of tumor cells, antibody-dependent cellular cytotoxicity and natural killing; and iv) T-lymphocytes possessing antigen-specific receptors and having the capacity to recognize a tumor cell carrying complementary marker molecules (Schreiber, H., 1989, in Fundamental Immunology (ed). W. E. Paul, pp. 923-955).

Cancer immunotherapy generally focuses on inducing humoral immune responses, cellular immune responses, or both. Moreover, it is well established that induction of CD4⁺ T helper cells is necessary in order to secondarily induce either antibodies or cytotoxic CD8⁺ T cells. Polypeptide antigens that are selective or ideally
5 specific for cancer cells, particularly ovarian cancer cells, offer a powerful approach for inducing immune responses against ovarian cancer, and are an important aspect of the present invention.

Therefore, in further aspects of the present invention, the pharmaceutical compositions described herein may be used to stimulate an immune response against
10 cancer, particularly for the immunotherapy of ovarian cancer. Within such methods, the pharmaceutical compositions described herein are administered to a patient, typically a warm-blooded animal, preferably a human. A patient may or may not be afflicted with cancer. Pharmaceutical compositions and vaccines may be administered either prior to or following surgical removal of primary tumors and/or treatment such as
15 administration of radiotherapy or conventional chemotherapeutic drugs. As discussed above, administration of the pharmaceutical compositions may be by any suitable method, including administration by intravenous, intraperitoneal, intramuscular, subcutaneous, intranasal, intradermal, anal, vaginal, topical and oral routes.

Within certain embodiments, immunotherapy may be active
20 immunotherapy, in which treatment relies on the *in vivo* stimulation of the endogenous host immune system to react against tumors with the administration of immune response-modifying agents (such as polypeptides and polynucleotides as provided herein).

Within other embodiments, immunotherapy may be passive
25 immunotherapy, in which treatment involves the delivery of agents with established tumor-immune reactivity (such as effector cells or antibodies) that can directly or indirectly mediate antitumor effects and does not necessarily depend on an intact host immune system. Examples of effector cells include T cells as discussed above, T lymphocytes (such as CD8⁺ cytotoxic T lymphocytes and CD4⁺ T-helper tumor-
30 infiltrating lymphocytes), killer cells (such as Natural Killer cells and lymphokine-activated killer cells), B cells and antigen-presenting cells (such as dendritic cells and

macrophages) expressing a polypeptide provided herein. T cell receptors and antibody receptors specific for the polypeptides recited herein may be cloned, expressed and transferred into other vectors or effector cells for adoptive immunotherapy. The polypeptides provided herein may also be used to generate antibodies or anti-idiotypic
5 antibodies (as described above and in U.S. Patent No. 4,918,164) for passive immunotherapy.

Monoclonal antibodies may be labeled with any of a variety of labels for desired selective usages in detection, diagnostic assays or therapeutic applications (as described in U.S. Patent Nos. 6,090,365; 6,015,542; 5,843,398; 5,595,721; and
10 4,708,930, hereby incorporated by reference in their entirety as if each was incorporated individually). In each case, the binding of the labelled monoclonal antibody to the determinant site of the antigen will signal detection or delivery of a particular therapeutic agent to the antigenic determinant on the non-normal cell. A further object of this invention is to provide the specific monoclonal antibody suitably labelled for
15 achieving such desired selective usages thereof.

Effector cells may generally be obtained in sufficient quantities for adoptive immunotherapy by growth *in vitro*, as described herein. Culture conditions for expanding single antigen-specific effector cells to several billion in number with retention of antigen recognition *in vivo* are well known in the art. Such *in vitro* culture
20 conditions typically use intermittent stimulation with antigen, often in the presence of cytokines (such as IL-2) and non-dividing feeder cells. As noted above, immunoreactive polypeptides as provided herein may be used to rapidly expand antigen-specific T cell cultures in order to generate a sufficient number of cells for immunotherapy. In particular, antigen-presenting cells, such as dendritic, macrophage,
25 monocyte, fibroblast and/or B cells, may be pulsed with immunoreactive polypeptides or transfected with one or more polynucleotides using standard techniques well known in the art. For example, antigen-presenting cells can be transfected with a polynucleotide having a promoter appropriate for increasing expression in a recombinant virus or other expression system. Cultured effector cells for use in therapy
30 must be able to grow and distribute widely, and to survive long term *in vivo*. Studies have shown that cultured effector cells can be induced to grow *in vivo* and to survive

long term in substantial numbers by repeated stimulation with antigen supplemented with IL-2 (*see, for example, Cheever et al., Immunological Reviews 157:177, 1997*).

Alternatively, a vector expressing a polypeptide recited herein may be introduced into antigen presenting cells taken from a patient and clonally propagated *ex vivo* for transplant back into the same patient. Transfected cells may be reintroduced into the patient using any means known in the art, preferably in sterile form by intravenous, intracavitary, intraperitoneal or intratumor administration.

Routes and frequency of administration of the therapeutic compositions described herein, as well as dosage, will vary from individual to individual, and may be readily established using standard techniques. In general, the pharmaceutical compositions and vaccines may be administered by injection (*e.g.*, intracutaneous, intramuscular, intravenous or subcutaneous), intranasally (*e.g.*, by aspiration) or orally. Preferably, between 1 and 10 doses may be administered over a 52 week period. Preferably, 6 doses are administered, at intervals of 1 month, and booster vaccinations may be given periodically thereafter. Alternate protocols may be appropriate for individual patients. A suitable dose is an amount of a compound that, when administered as described above, is capable of promoting an anti-tumor immune response, and is at least 10-50% above the basal (*i.e.*, untreated) level. Such response can be monitored by measuring the anti-tumor antibodies in a patient or by vaccine-dependent generation of cytolytic effector cells capable of killing the patient's tumor cells *in vitro*. Such vaccines should also be capable of causing an immune response that leads to an improved clinical outcome (*e.g.*, more frequent remissions, complete or partial or longer disease-free survival) in vaccinated patients as compared to non-vaccinated patients. In general, for pharmaceutical compositions and vaccines comprising one or more polypeptides, the amount of each polypeptide present in a dose ranges from about 25 μ g to 5 mg per kg of host. Suitable dose sizes will vary with the size of the patient, but will typically range from about 0.1 mL to about 5 mL.

In general, an appropriate dosage and treatment regimen provides the active compound(s) in an amount sufficient to provide therapeutic and/or prophylactic benefit. Such a response can be monitored by establishing an improved clinical outcome (*e.g.*, more frequent remissions, complete or partial, or longer disease-free

survival) in treated patients as compared to non-treated patients. Increases in preexisting immune responses to a tumor protein generally correlate with an improved clinical outcome. Such immune responses may generally be evaluated using standard proliferation, cytotoxicity or cytokine assays, which may be performed using samples
5 obtained from a patient before and after treatment.

Cancer Detection and Diagnostic Compositions, Methods and Kits

In general, a cancer may be detected in a patient based on the presence of one or more ovarian tumor proteins and/or polynucleotides encoding such proteins in a biological sample (for example, blood, sera, sputum urine and/or tumor biopsies)
10 obtained from the patient. In other words, such proteins may be used as markers to indicate the presence or absence of a cancer such as ovarian cancer. In addition, such proteins may be useful for the detection of other cancers. The binding agents provided herein generally permit detection of the level of antigen that binds to the agent in the biological sample.

15 Polynucleotide primers and probes may be used to detect the level of mRNA encoding a tumor protein, which is also indicative of the presence or absence of a cancer. In general, a tumor sequence should be present at a level that is at least two-fold, preferably three-fold, and more preferably five-fold or higher in tumor tissue than in normal tissue of the same type from which the tumor arose. Expression levels of a
20 particular tumor sequence in tissue types different from that in which the tumor arose are irrelevant in certain diagnostic embodiments since the presence of tumor cells can be confirmed by observation of predetermined differential expression levels, e.g., 2-fold, 5-fold, etc, in tumor tissue to expression levels in normal tissue of the same type.

Other differential expression patterns can be utilized advantageously for
25 diagnostic purposes. For example, in one aspect of the invention, overexpression of a tumor sequence in tumor tissue and normal tissue of the same type, but not in other normal tissue types, e.g. PBMCs, can be exploited diagnostically. In this case, the presence of metastatic tumor cells, for example in a sample taken from the circulation or some other tissue site different from that in which the tumor arose, can be identified
30 and/or confirmed by detecting expression of the tumor sequence in the sample, for

example using RT-PCR analysis. In many instances, it will be desired to enrich for tumor cells in the sample of interest, e.g., PBMCs, using cell capture or other like techniques.

There are a variety of assay formats known to those of ordinary skill in the art for using a binding agent to detect polypeptide markers in a sample. See, e.g., Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988. In general, the presence or absence of a cancer in a patient may be determined by (a) contacting a biological sample obtained from a patient with a binding agent; (b) detecting in the sample a level of polypeptide that binds to the binding agent; and (c) comparing the level of polypeptide with a predetermined cut-off value.

In a preferred embodiment, the assay involves the use of binding agent immobilized on a solid support to bind to and remove the polypeptide from the remainder of the sample. The bound polypeptide may then be detected using a detection reagent that contains a reporter group and specifically binds to the binding agent/polypeptide complex. Such detection reagents may comprise, for example, a binding agent that specifically binds to the polypeptide or an antibody or other agent that specifically binds to the binding agent, such as an anti-immunoglobulin, protein G, protein A or a lectin. Alternatively, a competitive assay may be utilized, in which a polypeptide is labeled with a reporter group and allowed to bind to the immobilized binding agent after incubation of the binding agent with the sample. The extent to which components of the sample inhibit the binding of the labeled polypeptide to the binding agent is indicative of the reactivity of the sample with the immobilized binding agent. Suitable polypeptides for use within such assays include full length ovarian tumor proteins and polypeptide portions thereof to which the binding agent binds, as described above.

The solid support may be any material known to those of ordinary skill in the art to which the tumor protein may be attached. For example, the solid support may be a test well in a microtiter plate or a nitrocellulose or other suitable membrane. Alternatively, the support may be a bead or disc, such as glass, fiberglass, latex or a plastic material such as polystyrene or polyvinylchloride. The support may also be a magnetic particle or a fiber optic sensor, such as those disclosed, for example, in U.S.

Patent No. 5,359,681. The binding agent may be immobilized on the solid support using a variety of techniques known to those of skill in the art, which are amply described in the patent and scientific literature. In the context of the present invention, the term "immobilization" refers to both noncovalent association, such as adsorption, and covalent attachment (which may be a direct linkage between the agent and functional groups on the support or may be a linkage by way of a cross-linking agent). Immobilization by adsorption to a well in a microtiter plate or to a membrane is preferred. In such cases, adsorption may be achieved by contacting the binding agent, in a suitable buffer, with the solid support for a suitable amount of time. The contact time varies with temperature, but is typically between about 1 hour and about 1 day. In general, contacting a well of a plastic microtiter plate (such as polystyrene or polyvinylchloride) with an amount of binding agent ranging from about 10 ng to about 10 μ g, and preferably about 100 ng to about 1 μ g, is sufficient to immobilize an adequate amount of binding agent.

Covalent attachment of binding agent to a solid support may generally be achieved by first reacting the support with a bifunctional reagent that will react with both the support and a functional group, such as a hydroxyl or amino group, on the binding agent. For example, the binding agent may be covalently attached to supports having an appropriate polymer coating using benzoquinone or by condensation of an aldehyde group on the support with an amine and an active hydrogen on the binding partner (*see, e.g.,* Pierce Immunotechnology Catalog and Handbook, 1991, at A12-A13).

In certain embodiments, the assay is a two-antibody sandwich assay. This assay may be performed by first contacting an antibody that has been immobilized on a solid support, commonly the well of a microtiter plate, with the sample, such that polypeptides within the sample are allowed to bind to the immobilized antibody. Unbound sample is then removed from the immobilized polypeptide-antibody complexes and a detection reagent (preferably a second antibody capable of binding to a different site on the polypeptide) containing a reporter group is added. The amount of detection reagent that remains bound to the solid support is then determined using a method appropriate for the specific reporter group.

More specifically, once the antibody is immobilized on the support as described above, the remaining protein binding sites on the support are typically blocked. Any suitable blocking agent known to those of ordinary skill in the art, such as bovine serum albumin or Tween 20™ (Sigma Chemical Co., St. Louis, MO). The
5 immobilized antibody is then incubated with the sample, and polypeptide is allowed to bind to the antibody. The sample may be diluted with a suitable diluent, such as phosphate-buffered saline (PBS) prior to incubation. In general, an appropriate contact time (*i.e.*, incubation time) is a period of time that is sufficient to detect the presence of polypeptide within a sample obtained from an individual with ovarian least about 95%
10 of that achieved at equilibrium between bound and unbound polypeptide. Those of ordinary skill in the art will recognize that the time necessary to achieve equilibrium may be readily determined by assaying the level of binding that occurs over a period of time. At room temperature, an incubation time of about 30 minutes is generally sufficient.

15 Unbound sample may then be removed by washing the solid support with an appropriate buffer, such as PBS containing 0.1% Tween 20™. The second antibody, which contains a reporter group, may then be added to the solid support. Preferred reporter groups include those groups recited above.

The detection reagent is then incubated with the immobilized antibody-polypeptide complex for an amount of time sufficient to detect the bound polypeptide.
20 An appropriate amount of time may generally be determined by assaying the level of binding that occurs over a period of time. Unbound detection reagent is then removed and bound detection reagent is detected using the reporter group. The method employed for detecting the reporter group depends upon the nature of the reporter group. For
25 radioactive groups, scintillation counting or autoradiographic methods are generally appropriate. Spectroscopic methods may be used to detect dyes, luminescent groups and fluorescent groups. Biotin may be detected using avidin, coupled to a different reporter group (commonly a radioactive or fluorescent group or an enzyme). Enzyme reporter groups may generally be detected by the addition of substrate (generally for a
30 specific period of time), followed by spectroscopic or other analysis of the reaction products.

To determine the presence or absence of a cancer, such as ovarian cancer, the signal detected from the reporter group that remains bound to the solid support is generally compared to a signal that corresponds to a predetermined cut-off value. In one preferred embodiment, the cut-off value for the detection of a cancer is the average mean signal obtained when the immobilized antibody is incubated with samples from patients without the cancer. In general, a sample generating a signal that is three standard deviations above the predetermined cut-off value is considered positive for the cancer. In an alternate preferred embodiment, the cut-off value is determined using a Receiver Operator Curve, according to the method of Sackett et al., *Clinical Epidemiology: A Basic Science for Clinical Medicine*, Little Brown and Co., 1985, p. 106-7. Briefly, in this embodiment, the cut-off value may be determined from a plot of pairs of true positive rates (*i.e.*, sensitivity) and false positive rates (100%-specificity) that correspond to each possible cut-off value for the diagnostic test result. The cut-off value on the plot that is the closest to the upper left-hand corner (*i.e.*, the value that encloses the largest area) is the most accurate cut-off value, and a sample generating a signal that is higher than the cut-off value determined by this method may be considered positive. Alternatively, the cut-off value may be shifted to the left along the plot, to minimize the false positive rate, or to the right, to minimize the false negative rate. In general, a sample generating a signal that is higher than the cut-off value determined by this method is considered positive for a cancer.

In a related embodiment, the assay is performed in a flow-through or strip test format, wherein the binding agent is immobilized on a membrane, such as nitrocellulose. In the flow-through test, polypeptides within the sample bind to the immobilized binding agent as the sample passes through the membrane. A second, labeled binding agent then binds to the binding agent-polypeptide complex as a solution containing the second binding agent flows through the membrane. The detection of bound second binding agent may then be performed as described above. In the strip test format, one end of the membrane to which binding agent is bound is immersed in a solution containing the sample. The sample migrates along the membrane through a region containing second binding agent and to the area of immobilized binding agent. Concentration of second binding agent at the area of immobilized antibody indicates the

presence of a cancer. Typically, the concentration of second binding agent at that site generates a pattern, such as a line, that can be read visually. The absence of such a pattern indicates a negative result. In general, the amount of binding agent immobilized on the membrane is selected to generate a visually discernible pattern when the
5 biological sample contains a level of polypeptide that would be sufficient to generate a positive signal in the two-antibody sandwich assay, in the format discussed above. Preferred binding agents for use in such assays are antibodies and antigen-binding fragments thereof. Preferably, the amount of antibody immobilized on the membrane ranges from about 25 ng to about 1 μ g, and more preferably from about 50 ng to about
10 500 ng. Such tests can typically be performed with a very small amount of biological sample.

Of course, numerous other assay protocols exist that are suitable for use with the tumor proteins or binding agents of the present invention. The above descriptions are intended to be exemplary only. For example, it will be apparent to
15 those of ordinary skill in the art that the above protocols may be readily modified to use tumor polypeptides to detect antibodies that bind to such polypeptides in a biological sample. The detection of such tumor protein specific antibodies may correlate with the presence of a cancer.

A cancer may also, or alternatively, be detected based on the presence of
20 T cells that specifically react with a tumor protein in a biological sample. Within certain methods, a biological sample comprising CD4⁺ and/or CD8⁺ T cells isolated from a patient is incubated with a tumor polypeptide, a polynucleotide encoding such a polypeptide and/or an APC that expresses at least an immunogenic portion of such a polypeptide, and the presence or absence of specific activation of the T cells is detected.
25 Suitable biological samples include, but are not limited to, isolated T cells. For example, T cells may be isolated from a patient by routine techniques (such as by Ficoll/Hypaque density gradient centrifugation of peripheral blood lymphocytes). T cells may be incubated *in vitro* for 2-9 days (typically 4 days) at 37°C with polypeptide (e.g., 5 - 25 μ g/ml). It may be desirable to incubate another aliquot of a T cell sample in
30 the absence of tumor polypeptide to serve as a control. For CD4⁺ T cells, activation is preferably detected by evaluating proliferation of the T cells. For CD8⁺ T cells,

activation is preferably detected by evaluating cytolytic activity. A level of proliferation that is at least two fold greater and/or a level of cytolytic activity that is at least 20% greater than in disease-free patients indicates the presence of a cancer in the patient.

As noted above, a cancer may also, or alternatively, be detected based on
5 the level of mRNA encoding a tumor protein in a biological sample. For example, at least two oligonucleotide primers may be employed in a polymerase chain reaction (PCR) based assay to amplify a portion of a tumor cDNA derived from a biological sample, wherein at least one of the oligonucleotide primers is specific for (*i.e.*, hybridizes to) a polynucleotide encoding the tumor protein. The amplified cDNA is
10 then separated and detected using techniques well known in the art, such as gel electrophoresis.

Similarly, oligonucleotide probes that specifically hybridize to a polynucleotide encoding a tumor protein may be used in a hybridization assay to detect the presence of polynucleotide encoding the tumor protein in a biological sample.

15 To permit hybridization under assay conditions, oligonucleotide primers and probes should comprise an oligonucleotide sequence that has at least about 60%, preferably at least about 75% and more preferably at least about 90%, identity to a portion of a polynucleotide encoding a tumor protein of the invention that is at least 10 nucleotides, and preferably at least 20 nucleotides, in length. Preferably,
20 oligonucleotide primers and/or probes hybridize to a polynucleotide encoding a polypeptide described herein under moderately stringent conditions, as defined above. Oligonucleotide primers and/or probes which may be usefully employed in the diagnostic methods described herein preferably are at least 10-40 nucleotides in length. In a preferred embodiment, the oligonucleotide primers comprise at least 10 contiguous
25 nucleotides, more preferably at least 15 contiguous nucleotides, of a DNA molecule having a sequence as disclosed herein. Techniques for both PCR based assays and hybridization assays are well known in the art (*see*, for example, Mullis et al., *Cold Spring Harbor Symp. Quant. Biol.*, 51:263, 1987; Erlich ed., *PCR Technology*, Stockton Press, NY, 1989).

30 One preferred assay employs RT-PCR, in which PCR is applied in conjunction with reverse transcription. Typically, RNA is extracted from a biological

sample, such as biopsy tissue, and is reverse transcribed to produce cDNA molecules. PCR amplification using at least one specific primer generates a cDNA molecule, which may be separated and visualized using, for example, gel electrophoresis. Amplification may be performed on biological samples taken from a test patient and from an individual who is not afflicted with a cancer. The amplification reaction may be performed on several dilutions of cDNA spanning two orders of magnitude. A two-fold or greater increase in expression in several dilutions of the test patient sample as compared to the same dilutions of the non-cancerous sample is typically considered positive.

10 In another aspect of the present invention, cell capture technologies may be used in conjunction, with, for example, real-time PCR to provide a more sensitive tool for detection of metastatic cells expressing ovarian tumor antigens. Detection of ovarian cancer cells in biological samples, e.g., bone marrow samples, peripheral blood, and small needle aspiration samples is desirable for diagnosis and prognosis in ovarian cancer patients.

Immunomagnetic beads coated with specific monoclonal antibodies to surface cell markers, or tetrameric antibody complexes, may be used to first enrich or positively select cancer cells in a sample. Various commercially available kits may be used, including Dynabeads® Epithelial Enrich (DynaL Biotech, Oslo, Norway), StemSep™ (StemCell Technologies, Inc., Vancouver, BC), and RosetteSep (StemCell Technologies). A skilled artisan will recognize that other methodologies and kits may also be used to enrich or positively select desired cell populations. Dynabeads® Epithelial Enrich contains magnetic beads coated with mAbs specific for two glycoprotein membrane antigens expressed on normal and neoplastic epithelial tissues.

25 The coated beads may be added to a sample and the sample then applied to a magnet, thereby capturing the cells bound to the beads. The unwanted cells are washed away and the magnetically isolated cells eluted from the beads and used in further analyses.

RosetteSep can be used to enrich cells directly from a blood sample and consists of a cocktail of tetrameric antibodies that targets a variety of unwanted cells and crosslinks them to glycophorin A on red blood cells (RBC) present in the sample, forming rosettes. When centrifuged over Ficoll, targeted cells pellet along with the free

RBC. The combination of antibodies in the depletion cocktail determines which cells will be removed and consequently which cells will be recovered. Antibodies that are available include, but are not limited to: CD2, CD3, CD4, CD5, CD8, CD10, CD11b, CD14, CD15, CD16, CD19, CD20, CD24, CD25, CD29, CD33, CD34, CD36, CD38, 5 CD41, CD45, CD45RA, CD45RO, CD56, CD66B, CD66e, HLA-DR, IgE, and TCR $\alpha\beta$.

Additionally, it is contemplated in the present invention that mAbs specific for ovarian tumor antigens can be generated and used in a similar manner. For example, mAbs that bind to tumor-specific cell surface antigens may be conjugated to magnetic beads, or formulated in a tetrameric antibody complex, and used to enrich or 10 positively select metastatic ovarian tumor cells from a sample. Once a sample is enriched or positively selected, cells may be lysed and RNA isolated. RNA may then be subjected to RT-PCR analysis using ovarian tumor-specific primers in a real-time PCR assay as described herein. One skilled in the art will recognize that enriched or selected populations of cells may be analyzed by other methods (*e.g. in situ* hybridization or 15 flow cytometry).

In another embodiment, the compositions described herein may be used as markers for the progression of cancer. In this embodiment, assays as described above for the diagnosis of a cancer may be performed over time, and the change in the level of reactive polypeptide(s) or polynucleotide(s) evaluated. For example, the assays may be 20 performed every 24-72 hours for a period of 6 months to 1 year, and thereafter performed as needed. In general, a cancer is progressing in those patients in whom the level of polypeptide or polynucleotide detected increases over time. In contrast, the cancer is not progressing when the level of reactive polypeptide or polynucleotide either remains constant or decreases with time.

25 Certain *in vivo* diagnostic assays may be performed directly on a tumor. One such assay involves contacting tumor cells with a binding agent. The bound binding agent may then be detected directly or indirectly via a reporter group. Such binding agents may also be used in histological applications. Alternatively, polynucleotide probes may be used within such applications.

30 As noted above, to improve sensitivity, multiple tumor protein markers may be assayed within a given sample. It will be apparent that binding agents specific

for different proteins provided herein may be combined within a single assay. Further, multiple primers or probes may be used concurrently. The selection of tumor protein markers may be based on routine experiments to determine combinations that results in optimal sensitivity. In addition, or alternatively, assays for tumor proteins provided
5 herein may be combined with assays for other known tumor antigens.

The present invention further provides kits for use within any of the above diagnostic methods. Such kits typically comprise two or more components necessary for performing a diagnostic assay. Components may be compounds, reagents, containers and/or equipment. For example, one container within a kit may contain a
10 monoclonal antibody or fragment thereof that specifically binds to a tumor protein. Such antibodies or fragments may be provided attached to a support material, as described above. One or more additional containers may enclose elements, such as reagents or buffers, to be used in the assay. Such kits may also, or alternatively, contain a detection reagent as described above that contains a reporter group suitable for direct
15 or indirect detection of antibody binding.

Alternatively, a kit may be designed to detect the level of mRNA encoding a tumor protein in a biological sample. Such kits generally comprise at least one oligonucleotide probe or primer, as described above, that hybridizes to a polynucleotide encoding a tumor protein. Such an oligonucleotide may be used, for
20 example, within a PCR or hybridization assay. Additional components that may be present within such kits include a second oligonucleotide and/or a diagnostic reagent or container to facilitate the detection of a polynucleotide encoding a tumor protein.

The following Examples are offered by way of illustration and not by way of limitation.

25

EXAMPLES

EXAMPLE 1

IDENTIFICATION OF OVARIAN TUMOR PROTEIN CDNAS

30

This Example illustrates the identification of cDNA molecules encoding ovarian tumor proteins.

Ovarian-specific genes were identified by electronic subtraction. The method used was similar to that described by Vasmatizis et al., *Proc. Natl. Acad. Sci. USA* 95:300-304, 1998, but there were several key differences. The sequences of EST clones (1,453,679) were downloaded from the GenBank public human EST database.

5 Human cDNA libraries were downloaded to create a database of these cDNA libraries and the EST sequences derived from them. The cDNA libraries were grouped into three groups: Plus, Minus and Other/Neutral. The Plus group included 26 libraries constructed from ovary normal and tumor tissues (and therefore including those containing ovary-specific ESTs); the Minus group consisted of 184 libraries derived

10 from all adult normal tissues except for breast, uterus and male-specific tissues, such as prostate; the Other/Neutral group contained libraries from tissues where expression is considered irrelevant (e.g., fetal tissue, tumors, cell lines other than ovary cell lines, non-relevant tissues. A total of 43,581 ESTs were derived from the 26 available ovary libraries (this subset represents 3% of the total 1,453,679 ESTs in the database). These

15 ESTs were preprocessed to remove common sequence repeats and cloning adapters, resulting in a final Plus group of 39,019 (a decrease of 10%).

Each Plus group (ovary) EST sequence was used as a query "seed" sequence in a BLASTN (version 2.0.9; May 7, 1999) search against the total human EST database. Standard measures of similarity are insufficient in this sort of analysis,

20 as EST relationships often include short stretches and poor sequence data. Criteria employed in this study required a matching segment to be at least 75 nucleotides in length, and the density of exact matches within this segment to be at least 80%. This was considered conservative criteria designed to avoid short spurious matches while allowing for polymorphisms and errors in sequencing. Each BLAST search generated a

25 cluster of related sequences based on direct overlap with the query "seed" sequence. A second level of clustering was performed to merge closely related clusters and to eliminate redundancy resulting from the fact that similar clusters are generated if the clusters contain more than one seed (i.e., sequences from the Plus EST group). The resulting "super clusters" were discarded if they grew in size to 200 or more ESTs, since

30 these probably represented repetitive elements that were not removed by the initial

preprocessing of the seeds, or highly expressed genes such as those for ribosomal proteins. Superclusters were merged if they shared at least one third of their sequences.

The BLAST searches gave rise to a total of 38,979 clusters. In the first super clustering stage, 10,307 clusters grew beyond the limit of 200 clones. Following this step and the elimination of duplicates, there were 15,758 super clusters remaining. This number was reduced to 15,244 after adjacent clusters were merged. Resulting super clusters were analyzed to determine the tissue source of each EST clone contained within it and this expression profile was used to classify the superclusters into four groups: Type 1 – this supercluster contains EST clones found in the Plus group only, with no expression in the Minus or Other/Neutral group libraries; Type 2 – EST clones in the supercluster are found in the Plus and Other/Neutral group libraries, with no expression in the Minus group; Type 3 – super cluster EST clones found in all groups, but the number of ESTs in the Plus group is higher than in either of the Minus or Other/Neutral groups; Type 4 – super cluster EST clones found in all groups, but the number in the Plus group is higher than in the Minus group with expression in the Other/Neutral group non relevant. Sequences derived from the Plus library group that were placed in Types 1, 2 and 3 superclusters are provided herein as SEQ ID NO:1-10,862. The electronic subtraction procedures identified these sequences as having significant differential expression in ovary tissue.

Sequences which showed good electronic subtraction profiles are described in more detail in Table 2.

Table 2: Electronic Subtraction Analysis

CloneID IMAGE#	Accession #	SuperClstr	Plus	Minus	Neutral	EST/ GenBank Annotation
595243	AA164216/ AA164215	AA164216	1	0	8	Novel
594484	AA180009/ AA164745	AA164850	2	0	27	Reverse transcriptase/ human endogenous retrovirus
843028	AA488541/ AA488406	AI567987	15	0	22	Megakaryocyte potentiating factor; mesothelin or CAK1
76058	T59567/ T59521	T59567	1	0	12	ALEX3, armadillo repeat containing protein
544863	AA075131/ AA075149	AA075131	1	0	0	Cathepsin L, cysteine protease (major excreted protein, MEP)
741139	AA402754/ AA402207	AA402207	1	0	3	Developmental protein eyes absent (Eab1)
545242	AA076182/ AA076085	AA180012	4	0	16	STAT1, signal transducer/ activator of transcription 1
75673	T58484	T58484	1	0	0	Genomic clone, chromosome 21
811600	AA458533/ AA454609	AA454609	1	0	13	Forkhead transcription factor HFH-4
739457	AA477250/ AA477249	AA402253	2	0	3	Novel
77262	T50256/ T50206	T50256	1	0	0	Vector, bacteriophage lambda lacZ
77257	T50212	T50212	1	0	0	Aspartate aminotransferase; genomic E.coli clones U/ K protein
74768	T47079	T47079	1	0	0	Genomic clone, chromosome 13q12.11-12.3
593444	AI733904/ AI732611/ AA160259/ AA160258	AA160258	1	0	0	Novel
811063	AI734202/ AI732823/ AA485619/ AA485451	AA485451	1	0	1	DNA-binding protein A (HUMBPA)

CloneID IMAGE#	Accession #	SuperClstr	Plus	Minus	Neutral	EST/ GenBank Annotation
593218	AI733902/ AI732609/ AA159613/ AA159743	AA159743	1	0	1	Novel
595232	AI732624/ AA173518	AA173518	1	0	0	Genomic clone, chromosome 12p11-37.2-54.4
75759	T58522	T58522	2	0	16	cDNA clone FLJ10748, weakly similar to ringcanal protein
714232	AA293601	AA293601	1	0	2	Novel
809858	AA464305/ AA455119	AI567829	2	0	17	COP9 complex subunit 7a
489076	AA057195/ AA057129	AA423796	1	0	5	Elastin microfibril interfase located protein (EMI); ketohexokinase
771328	AA476222/ AA476223	AA476222	2	0	6	MG81
755093	AI821682/ AI820932/ AA482660/ AA482508	AA482508	1	0	0	Neuronal apoptosis inhibitory protein; basic transn factor 2 p44
771053	AA430718/ AA427528	AA430718	1	0	2	Zinc Finger Protein 26;contains element THR, repetitive element
77250	T50141	T50141	7	2	1	Excisionase (xis) and integrase (int) bacteriophage genes
810109	AA465048/ AA464979	AA292245	2	0	9	cDNA clone DKFZp434K022
136984	AF074992/ R35849	AA401975	3	0	2	cDNA YH90A09
809776	AA454784/ AA454732	AA454732	3	0	19	IL-16, lymphocyte chemoattractant factor
593439	AA165668/ AA165632	AA165632	3	2	31	GAG polyprotein/ endogenous retrovirus sequences
76081	T59590	T59590	1	0	0	Novel
724886	AA404613/ AA291468	AA464208	5	0	6	contains PTR7 repetitive element; T cell receptor gamma V1?

CloneID IMAGE#	Accession #	SuperClstr	Plus	Minus	Neutral	EST/ GenBank Annotation
724962	AA404631/ AA291495	AA402723	15	0	38	Pregnancy-Associated endometrial alpha2 globulin; = HPP14
283151	N45230	AA284811	2	0	6	Genomic clones, alpha albumin and MER27 repetitive element
595363	AA164920/ AA164919	AA402031	3	1	30	Interferon-gamma induced small cytokine subfamily B #10
544548	AA075105/ AA074952	AI567987	15	0	22	Megakaryocyte potentiating factor; mesothelin or CAK1
595220	AA164851/ AA164850	AA164850	2	0	27	Genomic clones
545443	AA079191/ AA079190	AA167449	5	1	55	Genomic clone containing X (inactive)-specific transcript (XIST)
545185	AA076101	AA076101	1	0	0	Acidic ribosomal phosphoprotein P0
714080	AA284545/ AA284815	AA284815	1	0	0	Megakaryocyte potentiating factor; mesothelin or CAK1
595449	AA173739/ AA173383	AA173383	5	0	33	Neurotensin receptor; G protein- coupled receptor (GPR39)
	AA176693/ AA173996	AA173996	2	1	3	Pak1, p21-activated serine/ threonine-protein kinase (JNK path)
	AI821564/ AA284659/ AA284658	AA284658	1	0	1	Novel
	AA075034/ AA075033	AA169452	2	1	27	C. elegans homolog, CGI-19
	AA394090/ AA293773	AA293773	1	0	6	cDNA clone 23870
	T57605	T57605	1	0	0	Novel
	AA456327/ AA454681	AA480796	2	1	18	Genomic clone, chromosome 22
	AA164782/ AA164781	AA164644	3	0	15	Novel
	AA166898/ AA166756	AA166756	1	0	0	Mixed genomic clones, contains Alu repetitive element

CloneID IMAGE#	Accession #	SuperClstr	Plus	Minus	Neutral	EST/ GenBank Annotation
	AA075203/ AA075202	AA075202	1	0	1	E.coli genomic DNA; ycbS, a putative outer membrane protein
	AA405395/ AA402883	AA402723	15	0	38	Pregnancy-Associated endometrial alpha2 globulin; = HPP14
	AA172236/ AA172056	AA172056	1	0	4	Genomic clone, chromosome 22 (cat eye syndrome region)
	T57611/ T57557	T57557	1	0	0	Novel
	AA419214/ AA419229	AA173383	5	0	33	Neurotensin receptor; G protein- coupled receptor (GPR39)
	AA171974/ AA171755	AA171755	1	0	3	Genomic clone, chromosome 14
	AA076270/ AA076269	AA076269	1	0	0	Novel
	AA464689	AA464689	1	0	9	Novel
	AA456446/ AA454554	AA477508	3	0	38	Myo-inositol 1-phosphate synthase
	T60090	T60090	1	0	0	Genomic clone, chromosome 1
	AA075211/ AA075126	AA074946	2	0	0	Nucleoside diphosphate kinase (DR-nm23); Ndk3-like
	AA076228/ AA076227	AA076228	1	0	3	C. elegans homolog, CGI-62
	AA075310	AA075310	1	0	9	Sperm membrane protein; nucleoporin; Ran-binding protein 2
	T58551/ T58501	T58501	1	0	36	Genomic clone, chromosome 22q13.31-13.33
	T59551	T59551	1	0	0	Reverse transcriptase-like protein -rat
	AA464691	AA405746	2	0	25	Genomic clone, chromosome Xp22
	T52869/ T52868	T52868	1	0	0	Contains repetitive DNA
	AF147380/ T59906/ T59850	T59850	2	0	0	cDNA clone YB67A06, contains CER/ MER22 repetitive element

CloneID IMAGE#	Accession #	SuperClstr	Plus	Minus	Neutral	EST/ GenBank Annotation
2216507	AI653489	AI653489	1	0	1	gDNA 12p13.3 PAC RPCI5-940J5
2292394	AI648506	AI648506	2	0	54	PTK7 protein tyrosine kinase 7 (PTK7), mRNA
2216151	AI654462	AI654452	2	0	3	
981215	AA526423	AA526423	1	0	1	
1373880	AA828788	AA828788	1	0	0	gDNA chromosome 5 clone CTD-2318A17
1373112	AA837621	AA837621	1	0	4	
2002493	AI224233	AI223766	10	0	37	
1373281	AA827974	AA526119	2	0	0	gDNA clone 1018D12 20q11.1-11.22
2216154	AI654454	AI567994	2	0	8	VSGP/ F-spondin, complete cds
755378	AA410595	AA410247	3	0	1	Inhibitor of DNA binding 4, dom/ neg HLH protein (ID4),
769666	AA428329	AA291377	2	0	2	gDNA RP11-343B21 from 2
2216636	AI653555	AI653555	1	0	1	gDNA RP5-1185I7 from 7q11.23-q21
1455124	AA909822	AA909822	1	0	19	mRNA for GARS-AIRS-GART
1374123	AA828611	AA828611	1	0	0	
770437	AA427513	AA427513	2	0	4	gDNA chromosome 8 clone RP11-4K16
739389	AA476860	AA856964	3	1	26	TLS-associated Ser-Arg protein 1 (TASR1), mRNA
1985442	AI254625	AI254625	1	0	4	
981039	AA526017	AA526017	1	0	40	Smad1 mRNA
755111	AA411372	AA411372	1	0	15	mRNA for KIAA0326 gene
1965439	AI356599	AA862214	2	0	28	cDNA FLJ10943 fis, clone OVARC1001360
755436	AA419044	AA292435	3	0	23	cDNA FLJ10023 fis, clone HEMBA1000608
742144	AA405975	AA405727	2	0	0	
1373104	AA837606	AA828060	3	0	0	gDNA clone B331M8 map 4q25

CloneID IMAGE#	Accession #	SuperClstr	Plus	Minus	Neutral	EST/ GenBank Annotation
1373878	AA828778	AA828778	1	0	0	CREB binding protein (Rubinstein-Taybi syndrome)(CREBBP
1454766	AA908470	AA908470	1	0	10	gDNA clone RP11-83M17 from 7q31
1373719	AA828882	AA828882	2	0	0	
1374009	AA828703	AA828138	2	0	4	hepatocellular carcinoma antigen 519 (HCA519)
1373380	AA828138	AA828138	2	0	4	hepatocellular carcinoma antigen 519 (HCA519)
1373381	AA828148	AA828038	2	0	0	gDNA clone RP3-525L6 on 6p22.3-
1964546	AI284862	AI284862	1	0	1	gDNA clone RP11-215O15 on chromosome 9
1192366	AA650614	AA650614	2	0	0	
2260607	AI590617	AI590617	1	0	5	Staufen (Drosophila, RNA- binding protein) (STAU)
1154834	AA642166	AA642166	1	0	0	
981113	AA526205	AA526205	1	0	0	
2292558	AI648453	AI648453	1	0	0	
981376	AA525937	AA525937	1	0	0	
1374125	AA828612	AA828612	1	0	0	
1374171	AA828647	AA828647	1	0	0	eIF3 p40 subunit gene, exon 7
1965552	AI364864	AI364864	1	0	8	gDNA clone RP5-850O15 on 1p32.3-34.2
981310	AA525800	AA525800	1	0	0	
1373500	AA847986	AA829098	2	0	0	
1373375	AA828135	AA828135	2	1	10	zinc finger protein (ZFD25) mRNA
1190617	AA650078	AA650078	1	0	0	
2292555	AI648459	AI648459	1	0	9	mitosin mRNA
755357	AA410276	AA293518	2	0	4	mRNA for KIAA1335 protein
1373287	AA827977	AA827977	1	0	4	
1192390	AA650618	AA650614	2	0	0	

CloneID IMAGE#	Accession #	SuperClstr	Plus	Minus	Neutral	EST/ GenBank Annotation
2260515	AI590596	AI590596	3	0	4	gDNA chromosome 5 clone CTD-2019M3
981372	AA525935	AA525935	2	1	3	peroxisomal 70 kD memb protein mRNA
980932	AA525774	AA525774	1	0	0	gDNA clone B271E1 on 4q25
1374020	AA828698	AA828797	5	0	6	gDNA chromosome 20 clones 97 and 127
1373311	AA828039	AA828039	1	0	3	gDNA chromosome 21 segment HS21C010
1047907	AA773935	AA362728	2	0	11	cDNA: FLJ21661 fis, clone COL08837
1154937	AA627343	AA627343	1	0	0	
1373504	AA847988	AA847988	1	0	3	gDNA 12 BAC RP11-588G21
1963836	AI280530	AI280530	1	0	0	
981337	AA525918	AA525918	1	0	0	gDNA chromosome 21 segment HS21C046
1373879	AA828787	AA828787	2	0	1	
980966	AA525874	AA525874	1	0	0	gDNA BAC clone CTB-95A14 from 7p15-p21
2292391	AI648510	AI648510	1	0	0	SYBL1 gene, exons 6-8
2292584	AI648468	AI648468	1	0	6	mRNA for KIAA1442 protein, partial cds
981068	AA526095	AA525221	2	0	0	
1373285	AA827976	AA827976	2	0	2	
1373419	AA826312	AA826312	1	0	0	
1374172	AA828640	AA828640	2	0	0	
1373400	AA826292	AA826292	1	0	0	AMP-activated protein kinase (hAMPK)
1373695	AA828597	AA828597	1	0	1	gDNA BAC 55C20 on chromosome 6.
1373807	AA828690	AA828690	1	0	0	
1373415	AA826310	AA826310	1	0	0	
981359	AA525929	AA525929	2	0	13	gDNA, chromosome 21q, section 89/ 105

CloneID IMAGE#	Accession #	SuperClstr	Plus	Minus	Neutral	EST/ GenBank Annotation
1373309	AA828038	AA828038	2	0	0	gDNA clone RP3-525L6 on 6p22.3-
2260575	AI591141	AI591141	1	0	3	protein Tyr phosphatase, receptor type, M (PTPRM)
755715	AA496499	AA496499	2	0	2	platelet derived growth factor C (PDGFC)
809959	AA454835	AA430746	3	0	29	EWS protein/ E1A enhancer binding protein
1373509	AA848000	AA829138	3	0	0	
1374033	AA829114	AA828882	2	0	0	
1373080	AA837588	AA837588	1	0	4	gDNA 3 BAC RP11-91B3
1373691	AA828596	AA828596	1	0	15	gDNA clone 747H23 on 6q13-15
1964868	AI355523	AI355523	1	0	3	gDNA BAC clone GS1-67A24 from 7q21.q21.2
1190662	AA650084	AA650084	1	0	9	phospholipase C-beta-3 (PLCB3) gene
1154908	AA627330	AA627330	1	0	0	
1985131	AI242020	AI242020	1	0	3	Ca2+-activated potassium channel SK3
1985090	AI241990	AI241990	7	0	1	
1985420	AI254602	AI252922	4	0	1	
1985633	AI254657	AI254657	1	0	0	gDNA clone RP3-421H16 on Xp11.3-21.
1985494	AI254938	AI254938	2	0	1	gDNA chromosome 14 BAC R-131H24
1985584	AI251343	AI254785	2	0	1	gDNA GS1-304P7 on 1q25.1-31.1
1985071	AI241981	AI241981	1	0	0	mRNA KIAA0741 gene product (IF2)
1985548	AI251317	AA425733	4	0	25	gDNA on 7q22
1985096	AI241992	AI241992	1	0	0	gDNA clone RP13-206I21 on Xp11.3-21.1
1985539	AI251255	AI251255	1	0	0	
1985503	AI254963	AI254963	1	0	0	
1985111	AI242006	AI242006	2	0	2	CLNS1A gene, intron 1

CloneID IMAGE#	Accession #	SuperClstr	Plus	Minus	Neutral	EST/ GenBank Annotation
1985085	AI241987	AI241987	1	0	0	mRNA KIAA0741 gene product (IF2)
810330	AA464144	AA428222	2	0	3	cytochrome B561, HCYTO B561
980920	AA525758	AA525758	1	0	0	cDNA FLJ10145 fis, clone HEMBA1003322
1985162	AI254100	AI254100	2	0	4	Lens epithelium-derived growth factor
930200	AA502530	AA502530	1	0	0	
1985535	AI251253	AI251253	1	0	2	cDNA DKFZp586B0319
1985593	AI251369	AI251369	1	0	0	
2003171	AI250559	AI250559	1	0	6	NY-REN-41 antigen mRNA
810462	AA457130	AA457130	3	0	0	gDNA 8q21.3: RICK gene
1985079	AI241985	AI241985	1	0	0	mRNA KIAA0741 gene product (IF2)
769684	AA428757	AA425850	4	1	17	
1985127	AI242019	AI144024	7	0	3	
1985630	AI254644	AI254644	1	0	0	gDNA, chromosome 21q, section 69/ 105
1985576	AI251339	AI251339	1	0	1	
1985558	AI251322	AI251322	1	0	0	AF-4 gene, exons 2 to 7 and Alu repeats
1984612	AI255127	AI254106	2	0	2	
1373305	AA828036	AA828036	1	0	0	
2002838	AI224248	AI054405	4	0	14	
1984571	AI251722	AI251722	1	0	0	
1985520	AI251235	AI251235	2	0	3	GAC-1 (GAC-1) mRNA
1985272	AI254763	AI254763	1	0	0	
2002656	AI223675	AI252423	4	2	32	gDNA chromosome X region
1985594	AI251359	AI054387	8	0	4	
1985474	AI254929	AI254929	1	0	0	mRNA KIAA0741 gene product (IF2)
2119169	AI401777	AA363377	1	0	1	

CloneID IMAGE#	Accession #	SuperClstr	Plus	Minus	Neutral	EST/ GenBank Annotation
755480	AA410639	AA410639	1	0	2	
1985049	AI251638	AI251638	1	0	0	
2006525	AI265778	AI265778	1	0	35	cDNA FLJ11653 fis, clone HEMBA1004538
1984955	AI252857	AI252857	1	0	0	mRNA KIAA0741 gene product (IF2)
810467	AA457141	AA457141	1	0	0	
993046	AA570694	AA363055	1	0	4	hypothetical protein FLJ10099 (FLJ10099)
810530	AA464548	AA464548	3	2	25	Claudin 7 (CLDN7)
1985621	AI254651	AI254651	1	0	0	
2002637	AI223666	AI223666	1	0	2	
755865	AA496479	AA496479	1	0	4	gDNA chromosome 21, NF1- related locus
1985448	AI254906	AI254906	1	0	0	
2292434	AI648563	AI648563	2	0	0	gDNA 12p12-21.3-21.8
810498	AA457154	AA457154	1	0	3	gDNA clone RP11-379I13 on chromosome 10
2003068	AI223766	AI223766	10	0	37	
2006294	AI279388	AI279388	1	0	8	Integin, alpha 9 (ITGA9), mRNA
1985556	AI251321	AI251321	1	0	1	
1985494	AI254938	AI254938	2	0	1	gDNA chromosome 14 clone BAC R-131H24
1985040	AI251623	AI254571	3	0	1	
1984903	AI254586	AI053894	50	5	68	
2292460	AI648530	AI648530	1	0	1	Wilms tumor associated protein (WIT-1)
1984941	AI252841	AI252841	2	1	3	gDNA chromosome 14 clone BAC R-804L24
770747	AA454520	AA454520	1	0	0	
1985061	AI251642	AI251642	1	0	0	
2260925	AI591032	AI591032	1	0	4	
2006359	AI264852	AI264852	1	0	33	

CloneID IMAGE#	Accession #	SuperClstr	Plus	Minus	Neutral	EST/ GenBank Annotation
2292490	AI648545	AI648545	1	0	13	
770522	AA456762	AA456762	1	0	16	
2292472	AI648533	AI648533	1	0	0	
2002584	AI223506	AI223506	1	0	0	
2006508	AI265914	AI265914	1	0	4	
2002702	AI223706	AI223706	1	0	0	
2002948	AI223609	AI053964	5	0	13	
2002540	AI223470	AI223470	1	0	0	
2260893	AI591004	AI591004	1	0	2	
2003136	AI250523	AI250523	1	0	0	
2260680	AI590757	AI590757	1	0	1	
2219675	AI738841	AI738841	1	0	0	
2006285	AI279392	AI279392	1	0	70	
1984994	AI252969	AI252969	1	0	3	
2260809	AI608710	AI608710	1	0	0	
2260492	AI590577	AI590577	1	0	1	
1964644	AI287912	AI287912	1	0	0	
770961	AA428171	AA428171	1	0	1	
810067	AA455305	AA165584	8	7	55	
2107496	AI380369	AA362711	1	0	2	
2292567	AI648462	AI648462	1	0	1	
2260591	AI590610	AI590610	1	0	0	
2261022	AI591055	AI591055	1	0	2	
2292702	AI612846	AA167381	2	0	21	
2002858	AI224270	AA166620	3	0	17	
2006307	AI279408	AI279408	1	0	8	
1984931	AI252836	AI252836	1	0	35	
2260917	AI591022	AI591022	1	0	0	
770927	AA433865	AA433865	1	0	8	
2006292	AI279387	AI279387	1	0	2	

CloneID IMAGE#	Accession #	SuperClstr	Plus	Minus	Neutral	EST/ GenBank Annotation
2261010	AI609587	AI609587	2	0	14	
741039	AA402296	AA402296	1	0	0	
2006364	AI264845	AI264845	1	0	14	
1985021	AI251613	AI251613	1	0	0	
1374360	AA828959	AA829097	3	0	1	
1985005	AI252985	AI252985	1	0	1	
2258189	AI623135	AF136413	1	0	1	
2006320	AI279404	AI279404	1	0	6	
770715	AA454497	AA454497	1	0	9	
2261186	AI609337	AI609337	1	0	16	
2216400	AI654744	AI654744	1	0	16	
2006496	AI265908	AI265908	1	0	0	
2292705	AI612857	AI612857	1	0	0	
1964750	AI283331	AI283331	1	0	0	
2260699	AI590770	AI590770	1	0	9	
1984950	AI252845	AI254780	5	0	2	
770978	AA428247	AA428247	2	1	12	
770998	AA428253	AA430746	3	0	29	
2292390	AI648504	AI648504	1	0	1	
1985408	AI254597	AI254597	1	0	0	
2251282	AI658484	AA363270	1	0	1	
1373633	AA837669	AA837669	1	0	0	
2122708	AI521444	AI521444	1	0	33	
753401	AA406401	T29223	1	0	13	
2216215	AI654529	AI654529	1	0	0	
810889	AA459288	AA454869	5	3	76	
811573	AA458520	AA458520	1	0	1	
2216412	AI654749	AA464522	3	0	12	
754929	AA422036	AA422036	2	0	0	
981382	AA525940	AA525940	1	0	0	

CloneID IMAGE#	Accession #	SuperClstr	Plus	Minus	Neutral	EST/ GenBank Annotation
1965573	AJ364874	AJ364874	1	0	11	
981208	AA526408	AA526408	1	0	0	
757030	AA429021	AA402679	24	12	79	
742056	AA405746	AA405746	2	0	25	
1861875	AJ053808	AJ053808	1	0	17	
1373857	AA828779	AA482061	2	0	0	
1373888	AA828792	AA828792	1	0	0	
810216	AA463996	AA402679	24	12	79	
1374124	AA828604	AA828604	1	0	0	
981025	AA526011	AA526011	1	0	0	
1212476	AA643476	AA643450	4	0	8	
800175	AA581095	AA581095	1	0	0	
741705	AA402081	AA402081	1	0	0	
740320	AA477801	AA482651	2	0	6	
756537	AA436442	AA292747	2	0	2	
897358	AA496757	AA363048	1	0	11	
1965502	AJ364842	AJ364842	1	0	0	
755590	AA419204	AA419204	1	0	0	
1964045	AJ280624	AJ280624	1	0	0	
1965524	AJ364851	AJ364851	1	0	0	
740416	AA477828	AA422036	2	0	0	
770267	AA434256	AA434256	1	0	0	
811005	AA485364	AA292026	2	0	20	
741183	AA402548	AA402548	1	0	4	
727859	AA393566	AA362742	1	0	34	

EXAMPLE 2ANALYSIS OF CDNA EXPRESSION USING
MICROARRAY ANALYSIS AND REAL TIME PCR

The electronic subtraction performed in Example 1 identified 10,862
5 sequences. Of the original sequences, 1688 were excluded from continued analysis due to duplication of the sequences. The remaining 9174 sequences were evaluated for their expression profiles in a variety of tumor and normal tissue types using microarray analysis.

Using this approach, the cDNA sequences were PCR amplified using
10 vector specific primers. The resulting PCR products were then arrayed randomly onto one of two chips. Ovary Chip #3 contained 2784 EST clones and Ovary Chip #4 contained 4882 clones.

Their mRNA profiles in tumor and normal tissues were then examined using microarray technology essentially as described (Shena M. *et al.*, 1995 Science
15 270:467-470). The probes were selected from two groups, one of which is fluorescently labeled with Cy3 (Group I, typically tumor tissues, see Table I) and the other with Cy5 (Group II, typically normal tissue [N], see Table 3). The Chips were then hybridized with probe pairs (one from Group I and one from Group II) and the fluorescent intensities measured. The mean fluorescent intensity of probes from Group I was then
20 compared to the mean fluorescence of probes from Group II. A ratio was then determined using computational analysis. This ratio determines the relevance of a particular sequence to different tissues, both normal and tumor. In addition to this ratio, sequences were also classified based on the level of background present in Group II. If the mean fluorescent signal was <0.1 , this identified sequences with low background,
25 between 0.1-0.2 identified sequences with some background, and <0.2 identified sequences with potentially high background.

Ovary Chip #3 was probed with 35 different probe pairs (see Table 3, Groups I and II for details). Two different types of analysis were then performed:

1. The mean fluorescent intensity of ovary tumor tissue compared to the
30 mean fluorescent intensity of normal tissue (minus ovary). Sequences were selected using this analysis if the ratio was greater than three and the background <0.2 .

2. The mean fluorescent intensity of ovary tumor and ovary normal tissue compared to the mean fluorescent intensity of essential normal tissue (minus ovary). Sequences were selected using this analysis if the ratio was greater than 2.5 and the background <0.2 .

Table 3: Probes Used to Analyze Ovary Chip #3

Cy3 Probe		Cy5 Probe	
Ovary Tumor (IIIC)	261A	SPACT37	Adrenal Gland (essential)
Ovary Tumor (IIIC)	264A	396A	Skin N (essential)
Ovary Tumor (IIIC)	265A	SPACT56	Thymus N (essential)
Ovary Tumor (IIIC)	288A	600C	Bronchus N (essential)
Papillary Serous Cystic Neo. (IA)	854A	785B	Bronchus N (essential)
Papillary Serous (IA)	855A	407B	Bone N
Papillary Serous (IA)	856A	484A	Peritoneum Epithelial N (essential)
Papillary Serous Adeno (IA)	603A	SPACT52	Pituitary Gland N (essential)
Papillary Serous (IB)	857A	SPACT40	Skeletal muscle N (essential)
Papillary Serous Adeno (IB)	385A	SPACT55	Stomach (essential)
Papillary Serous Neo (IC)	492A	SPACT54	Spleen N (essential)
Papillary Serous (IIA)	858A	862A	Pancreas N (essential)
Papillary Serous (IIB)	859A	S27	Ovary N
Serous Borderline Tumor (IIIC)	605A	SPACT45	Spinal cord N (essential)
Papillary Serous Carcinoma (III)	495A	SPAAm1	Heart N
Mucinous Cystadeno (IB)	381C	S7	Ovary N
Mucinous Adeno (IIIC)	382A	S449A	Ovary N
Mets Mucinous Adeno (IIIC)	428B	SPACT53	Small intestine N (essential)
Endometrioid Adeno. (IA)	491A	502B	Esophagus N (essential)
Endometrioid Adeno. (II)	335A	199A	Colon N (essential)
Endometrioid Adeno. (II-III)	494A	SPACT46	Thyroid gland N (essential)
Cystic Endometrioid. (IIIC)	860A	783A	PBMC (resting) (essential)
Clear cell adeno. (I)	604A	415A	Aorta N (essential)

Clear cell (IA)	607A	776A	Trachea N (essential)
Granulosa Cell Tumor (IA)	S25	CT25	Trachea N (essential)
Granulosa Cell Tumor (IA)	S22	PAN2000	Pancreas N (pool) (essential)
Germ Cell Tumor	386A	S92	Breast (HMEC) N (essential)
Papillary Serous Adeno. (IIIB)	602A	328B/C	Bladder N (essential)
Papillary Serous Adeno. (IIIC)	S23	SPACT49	Bone marrow N (essential)
Papillary Serous Cystadeno. (IIIB)	606A	SPAAm2	Lung N (essential)
Mets Papillary Adeno. (IIIA)	383A	302B	Kidney N (essential)
Mets Papillary Serous. (IIIB)	384A	S44/782A	PBMC (activated) (essential)
Mets. Pap. Serous Cystadeno (IIIC)	426A	603A	Ovary tumor (Cy3 match)
Mets Pap. Adeno (IIIC)	429A	270B	Liver N (essential)
Ovary Tumor	427A	SPACT50	Brain N (essential)

Microarray analysis of clones that showed significant expression levels in tumor samples when compared to normals is presented in Table 4.

Table 4: Microarray Analysis From Ovary Chip #3

CloneID IMAGE#	Accession #	Ratio	Mean Signal 1	Mean Signal 2	Candidate Name
Threshold >3.0; Mean signal 2 <0.1; Ranked by Mean signal 1.					
595243	AA164216/AA164215	6.67	0.632	0.095	O1028C
594484	AA180009/AA164745	5.6	0.432	0.077	O1029C
843028	AA488541/AA488406	3.85	0.368	0.096	O1030C-R
76058	T59567/T59521	3.07	0.278	0.09	O1031C
544863	AA075131/AA075149	3.24	0.248	0.077	
741139	AA402754/AA402207	3.3	0.225	0.068	
545242	AA076182/AA076085	3.23	0.225	0.07	
75673	T58484	3.18	0.189	0.059	
811600	AA458533/AA454609	3.95	0.183	0.046	O1032C
739457	AA477250/AA477249	3.26	0.183	0.056	
77262	T50256/T50206	3.12	0.175	0.056	
77257	T50212	3.06	0.172	0.056	
74768	T47079	3.02	0.17	0.056	
593444	AI733904/AI732611/AA160259/ AA160258	4.96	0.164	0.033	
811063	AI734202/AI732823/AA485619/ AA485451	3.03	0.164	0.054	
593218	AI733902/AI732609/AA159613/ AA159743	3.01	0.149	0.05	
595232	AI732624/AA173518	4.19	0.144	0.034	
75759	T58522	3.68	0.142	0.038	
714232	AA293601	3.03	0.134	0.044	
809858	AA464305/AA455119	3.56	0.129	0.036	
489076	AA057195/AA057129	3.2	0.129	0.04	
771328	AA476222/AA476223	3.23	0.111	0.034	

CloneID IMAGE#	Accession #	Ratio	Mean Signal 1	Mean Signal 2	Candidate Name
755093	AI821682/AI820932/AA482660/ AA482508	5.95	0.094	0.016	
771053	AA430718/AA427528	3.37	0.086	0.025	
77250	T50141	3.22	0.082	0.025	
810109	AA465048/AA464979	4.65	0.078	0.017	
136984	AF074992/R35849	4.31	0.068	0.016	
809776	AA454784/AA454732	6.89	0.028	0.004	
Threshold >3.0; Signal 2 <0.2; Ranked by mean Signal 1.					
593439	AA165668/AA165632	9.75	1.149	0.118	O1033C
76081	T59590	3.86	0.708	0.183	
724886	AA404613/AA291468	3.04	0.532	0.175	
724962	AA404631/AA291495	3.36	0.438	0.13	HPP14
283151	N45230	3.2	0.379	0.119	
595363	AA164920/AA164919	3.04	0.31	0.102	
Threshold >3.0; Mean Signal 2 >0.2; Ranked by Mean Signal 1.					
544548	AA075105/AA074952	4.23	1.768	0.418	
595220	AA164851/AA164850	3.31	1.598	0.482	
545443	AA079191/AA079190	3.29	1.241	0.377	
545185	AA076101	3.04	1.152	0.379	
714080	AA284545/AA284815	5.09	1.05	0.206	
595449	AA173739/AA173383	4.95	0.999	0.202	O1034C
Threshold >2.5; Mean Signal 2<0.1; Ranked by Mean Signal 1.					
	AA176693/AA173996	15.64	1.43	0.091	
	AI821564/AA284659/ AA284658	4.71	0.323	0.069	
	AA075034/AA075033	3.01	0.291	0.097	O1035C
	AA394090/AA293773	3.04	0.275	0.091	O1036C

468

CloneID IMAGE#	Accession #	Ratio	Mean Signal 1	Mean Signal 2	Candidate Name
	T57605	2.92	0.19	0.065	
	AA456327/AA454681	3.19	0.188	0.059	
	AA164782/AA164781	2.85	0.138	0.049	
	AA166898/AA166756	2.73	0.129	0.047	
	AA075203/AA075202	3.35	0.128	0.038	
	AA405395/AA402883	3.07	0.126	0.041	
	AA172236/AA172056	3.17	0.124	0.039	
	T57611/T57557	2.62	0.118	0.045	
	AA419214/AA419229	2.68	0.112	0.042	
	AA171974/AA171755	2.89	0.109	0.038	
	AA076270/AA076269	5.32	0.087	0.016	
	AA464689	2.61	0.074	0.028	
	AA456446/AA454554	2.57	0.065	0.025	
	T60090	3.66	0.045	0.012	
	AA075211/AA075126	3.91	0.036	0.009	
Threshold >2.5; Mean Signal 2>2.0; Ranked by Mean Signal 1.					
	AA076228/AA076227	5.02	0.812	0.162	
	AA075310	4.98	0.582	0.117	
	T58551/T58501	3.15	0.528	0.167	
	T59551	3.13	0.431	0.138	
	AA464691	2.62	0.382	0.146	
	T52869/T52868	3.16	0.37	0.117	
	AF147380/T59906/T59850	2.55	0.335	0.131	
Threshold >2.5; Mean Signal 2 <0.1; Ranked by Mean Signal 1.					
	AA172293/AA171844	2.82	0.257	0.091	

CloneID IMAGE#	Accession #	Ratio	Mean Signal 1	Mean Signal 2	Candidate Name
	AA446928/AA443351	2.55	0.164	0.064	
	AA164691	3.09	0.107	0.035	
	AA160156/AA160155	2.68	0.088	0.033	
	AI732624/AA173518/AA17342 8	3.08	0.086	0.028	
595232	AI732624/AA173518	4.19	0.144	0.034	
	AA411128/AA292424	2.74	0.078	0.028	
	AI734109/AI732705/AA428413/ AA427400	2.55	0.072	0.028	

Of those sequences presented in Table 4, several were selected for Real Time PCR analysis of an extensive panel of normal and tumor tissues. Real Time PCR (see Gibson *et al.*, *Genome Research* 6:995-1001, 1996; Heid *et al.*, *Genome Research* 6:986-994, 1996) is a technique that evaluates the level of PCR product accumulation during amplification. This technique permits quantitative evaluation of mRNA levels in multiple samples. Briefly, this was done by extracting mRNA from tumor and normal tissue and preparing cDNA using standard techniques. Real Time PCR was then performed using a Perkin-Elmer/Applied Biosystems (Foster City, CA) 7700 Prism instrument. Matching primers and fluorescent probes were then designed for each of the 6 sequences and optimal concentrations for amplification and detection determined. The cDNA content was standardized using PCR specific for β -Actin. The amount of specific RNA was then determined by either comparing the number of copies present per 1000 picograms of β -Actin or by setting a baseline using the sample, which showed the lowest level of the gene of interest.

Table 5 contains the Real Time PCR results for 6 samples selected from Chip #3.

**Table 5: Real Time PCR Expression Profiles of Sequences Identified
on Ovary Chip #3**

Clone ID	Image #	SEQ ID NO	Ovary Tumor Tissue and tumor associated tissue	Normal Tissue
1029C	94484	1928; 10,864	HOE in 11/23; OE in 3/23; E 3/23; E in 2/2 SCID grown tumors; E in 1/1 tumor cell lines	OE in 1/3 Ovary
1030C	843028	10,196; 10,868; 10,870	HOE in 5/20; OE in 5/20; E in 9/20	E in ovary
1032C	811600	5085; 10,872; 10,873	HOE in 1/22; OE in 1/22; E in 14/22.	OE in Bronchia
1033C	593439	10,874; 10,875	HOE in 4/21; OE in 4/21; E 10/21	No expression detected
1034C	595449	2409; 10,876	HOE in 4/22; OE in 7/22; E in 7/22; OE in 1/2 SCID grown tumors.	E in brain, brain pool, pancreas,
O1036C	726384	3228; 3239; 10,880; 10,881; 10,882; 10883; 10,884	HOE in 1/13; OE in 4/13; E in 8/13; E in 2/2 SCID derived tissue.	E in PBMC (activated), Skeletal muscle, spleen, bone marrow, bronchia, brain

HOE= highly over expressed; OE= over expressed; E= expressed.

5 Ovary Chip #4 contained 4882 clones from the electronic subtraction process described in Example 1. This Chip was probed with 35 different probe paris

(see Table 6, Groups I and II for details). Two different types of analysis were then performed:

1. The mean fluorescent intensity of ovary tumor tissue compared to the mean fluorescent intensity of normal tissue (minus ovary). Sequences were selected using this analysis if the ratio was greater than three and the background <0.2.
2. The mean fluorescent intensity of ovary tumor and ovary normal tissue compared to the mean fluorescent intensity of essential normal tissue (minus ovary). Sequences were selected using this analysis if the ratio was greater than 2.5 and the background <0.2.

10

Table 6: Probes Used to Analyze Ovary Chip #4

Group I: Cy3 Probe		Group II: Cy5 Probe	
Ovary Tumor (IIIC)	201	SPACT84	Adrenal Gland N
Ovary Tumor (IIIC)	264A	SK2000	Skin Pool N
Ovary Tumor (IIIC)	263A	SPACT70	Thymus N
Ovary Tumor (IIIC)	288A	785A/D	Bronchus N
Papillary Serous adeno. (IA)	1068A	784B	Bronchus N
Serous Cystadenoma (IB)	1084A	741A/B	Bone N
Papillary Serous (IA)	856A	BP2000	Breast Pool N
Papillary Serous Adeno. (IA)	603A	SPACT67	Pituitary Gland N
Papillary Serous Cystadeno. (IB)	857A	SPACT75	Skeletal Muscle N
Papillary Serous Adeno. (IB)	385A	SPACT66	Stomach N
Papillary Serous Neo. (IC)	492A	SPACT69	Spleen N
Papillary Serous Cystadeno. (IIA)	858A	862A	Pancreas N
Mucinous Carcinoma (IIA)	1130A	S27	Ovary N
Mets. Papillary Serous (III)	495A	SPACT73	Heart N
Mets. Pap. Serous Cystadeno. (IIIB)	606A	S52/ SPACT78	Lung N
Papillary Serous Adeno. (IIIB)	602A	328B/C	Bladder N
Serous Borderline Tumor (IIIC)	605A	SPACT59	Spinal Cord N
Papillary Serous Adeno. (IIIC)	S23	SPACT49	Bone Marrow N
Mets. Papillary Adeno. (IIIA)	383A	KP2000	Kidney Pool N
Mets. Adenocarcinoma (IIIC)	1069A	1155A	PBMC (activated)
Mets. Papillary Adeno. (IIIC)	1070A	1070A	Ovary Tumor Cy3 match
Mets. Papillary Adeno. (III)	429A	SPACT81	Liver N
Mets. Pap. Serous Adeno. (IIIC)	427A	SPACT85	Brain N
Mucinous Cystadeno. (IB)	381C	S7	Ovary N
Mets. Endometrial Adeno.	1127A	S449A	Ovary N

(III)			
Mets. Mucinous Adeno. (IIIC)	428B	SPACT86	Small Intestine N
Endometrioid Carcinoma (IA)	1071A	502B	Esophagus N
Endometrioid Adeno. (IA)	1081A	1004A	Colon N
Mets. Adeno. (II-III)	494A	886A	Thyroid Gland N
Cystic Endometrioid Adeno. (IIIC)	860A	783A/888A	PBMC Pool (resting)
Clear Cell Adeno. (IA)	604A	415C	Aorta N
Clear Cell/Endometrioid (IA)	607A	CT25	Trachea Pool N
Granulosa Cell Tumor (IA)	S25	772B	Trachea N
Granulosa Cell Tumor (IA)	1082A	SPACT58	Pancreas Pool N
Ovary Germ-Cell Tumor (I)	386A	S12	Mammary Epithelial N

Microarray analysis of clones that showed significant expression levels in tumor samples when compared to normals is presented in Table 7.

5

Table 7: Microarray Analysis From Ovary Chip #4

CloneID IMAGE#	Accession #	Ratio	Mean Signal 1	Mean Signal 2	Candidate
Threshold >2.0; Mean Signal 2<0.1; Mean Signal 1 >0.2.					
		6.67	0.358	0.054	O1033C
		4.82	0.247	0.051	O1029C
		4.59	0.22	0.048	O1034C
		3.32	0.209	0.063	Phospholipase a2
2216507	AI653489	3.26	0.254	0.078	
2292394	AI648506	2.94	0.235	0.08	
2216151	AI654462	2.64	0.202	0.076	
981215	AA526423	2.56	0.227	0.089	
1373880	AA828788	2.16	0.201	0.093	
1373112	AA837621	2.1	0.201	0.096	
Threshold >2.0; Mean Signal 2 <0.1; Mean Signal 1=0.1-0.2.					
2002493	AI224233	15.14	0.102	0.007	

CloneID IMAGE#	Accession #	Ratio	Mean Signal 1	Mean Signal 2	Candidate
1373281	AA827974	4.19	0.102	0.024	
2216154	AI654454	3.52	0.195	0.055	
755378	AA410595	3.15	0.107	0.034	
769666	AA428329	2.97	0.112	0.038	
2216636	AI653555	2.92	0.123	0.042	
1455124	AA909822	2.89	0.183	0.063	
1374123	AA828611	2.86	0.176	0.061	
770437	AA427513	2.79	0.154	0.055	
739389	AA476860	2.78	0.153	0.055	
1985442	AI254625	2.72	0.111	0.041	
981039	AA526017	2.67	0.139	0.052	
755111	AA411372	2.66	0.11	0.042	
1965439	AI356599	2.65	0.136	0.051	
755436	AA419044	2.63	0.138	0.053	
742144	AA405975	2.58	0.104	0.04	
1373104	AA837606	2.56	0.172	0.067	
1373878	AA828778	2.55	0.156	0.061	
1454766	AA908470	2.55	0.152	0.06	
1373719	AA828882	2.41	0.163	0.068	
1374009	AA828703	2.36	0.184	0.078	
1373380	AA828138	2.34	0.17	0.073	
1373381	AA828148	2.32	0.161	0.07	
1964546	AI284862	2.29	0.1	0.044	
1192366	AA650614	2.27	0.157	0.069	
2260607	AI590617	2.25	0.102	0.045	
1154834	AA642166	2.25	0.155	0.069	
981113	AA526205	2.23	0.111	0.05	
2292558	AI648453	2.22	0.1	0.045	

CloneID IMAGE#	Accession #	Ratio	Mean Signal 1	Mean Signal 2	Candidate
981376	AA525937	2.21	0.138	0.063	
1374125	AA828612	2.21	0.159	0.072	
1374171	AA828647	2.21	0.183	0.082	
1965552	AI364864	2.21	0.142	0.064	
981310	AA525800	2.2	0.109	0.05	
1373500	AA847986	2.18	0.181	0.083	
1373375	AA828135	2.17	0.173	0.08	
1190617	AA650078	2.17	0.155	0.071	
2292555	AI648459	2.16	0.12	0.056	
755357	AA410276	2.15	0.153	0.071	
1373287	AA827977	2.15	0.166	0.077	
1192390	AA650618	2.14	0.182	0.085	
2260515	AI590596	2.13	0.18	0.085	
981372	AA525935	2.12	0.189	0.089	
980932	AA525774	2.11	0.199	0.094	
1374020	AA828698	2.11	0.178	0.085	
1373311	AA828039	2.09	0.157	0.075	
1047907	AA773935	2.09	0.158	0.076	
1154937	AA627343	2.09	0.197	0.094	
1373504	AA847988	2.08	0.19	0.091	
1963836	AI280530	2.08	0.187	0.09	
981337	AA525918	2.08	0.193	0.093	
1373879	AA828787	2.08	0.182	0.088	
980966	AA525874	2.07	0.199	0.096	
2292391	AI648510	2.07	0.115	0.056	
2292584	AI648468	2.06	0.104	0.051	
981068	AA526095	2.06	0.176	0.086	
1373285	AA827976	2.06	0.165	0.08	

475

CloneID IMAGE#	Accession #	Ratio	Mean Signal 1	Mean Signal 2	Candidate
1373419	AA826312	2.06	0.153	0.074	
1374172	AA828640	2.06	0.161	0.078	
1373400	AA826292	2.04	0.162	0.079	
OVM-84		2.04	0.166	0.081	OVM-84
1373695	AA828597	2.03	0.15	0.074	
1373807	AA828690	2.03	0.118	0.058	
1373415	AA826310	2.03	0.146	0.072	
981359	AA525929	2.03	0.172	0.084	
1373309	AA828038	2.03	0.16	0.079	
2260575	AI591141	2.02	0.107	0.053	
755715	AA496499	2.02	0.165	0.082	
809959	AA454835	2.02	0.152	0.075	
1373509	AA848000	2.01	0.193	0.096	
1374033	AA829114	2.01	0.172	0.086	
1373080	AA837588	2.01	0.187	0.093	
1373691	AA828596	2.01	0.193	0.096	
1964868	AI355523	2.01	0.109	0.054	
1190662	AA650084	2.01	0.179	0.089	
1154908	AA627330	2	0.178	0.089	
Threshold >2.0; Mean Signal 2 <0.1; Mean Signal 1 <0.1.					
1985131	AI242020	9.05	0.028	0.003	
1985090	AI241990	8.43	0.024	0.003	
1985420	AI254602	7.47	0.029	0.004	
1985633	AI254657	7.34	0.029	0.004	
1985494	AI254938	6.94	0.065	0.009	
1985584	AI251343	6.81	0.056	0.008	
1985071	AI241981	6.49	0.026	0.004	
1985548	AI251317	6.36	0.03	0.005	

CloneID IMAGE#	Accession #	Ratio	Mean Signal 1	Mean Signal 2	Candidate
1985096	AI241992	6.04	0.027	0.005	
1985539	AI251255	5.86	0.053	0.009	
1985503	AI254963	5.69	0.075	0.013	
1985111	AI242006	5.65	0.022	0.004	
1985085	AI241987	5.65	0.023	0.004	
810330	AA464144	5.52	0.018	0.003	
980920	AA525758	5.34	0.067	0.012	
1985162	AI254100	5.2	0.028	0.005	
930200	AA502530	5.17	0.035	0.007	
1985535	AI251253	5.09	0.044	0.009	
1985593	AI251369	5.07	0.06	0.012	
2003171	AI250559	4.92	0.026	0.005	
810462	AA457130	4.84	0.025	0.005	
1985079	AI241985	4.54	0.038	0.008	
769684	AA428757	4.42	0.019	0.004	
1985127	AI242019	4.42	0.031	0.007	
1985630	AI254644	4.41	0.059	0.013	
1985576	AI251339	4.37	0.042	0.01	
1985558	AI251322	4.36	0.074	0.017	
1984612	AI255127	4.35	0.04	0.009	
1373305	AA828036	4.32	0.08	0.018	
2002838	AI224248	4.06	0.015	0.004	
1984571	AI251722	3.92	0.028	0.007	
1985520	AI251235	3.78	0.045	0.012	
1985272	AI254763	3.75	0.038	0.01	
2002656	AI223675	3.62	0.022	0.006	
1985594	AI251359	3.6	0.039	0.011	
1985474	AI254929	3.47	0.053	0.015	

CloneID IMAGE#	Accession #	Ratio	Mean Signal 1	Mean Signal 2	Candidate
2119169	AI401777	3.11	0.07	0.023	
755480	AA410639	3.02	0.048	0.016	
1985049	AI251638	2.96	0.047	0.016	
2006525	AI265778	2.91	0.056	0.019	
1984955	AI252857	2.87	0.049	0.017	
810467	AA457141	2.83	0.071	0.025	
993046	AA570694	2.83	0.091	0.032	
810530	AA464548	2.77	0.018	0.007	
1985621	AI254651	2.73	0.037	0.014	
2002637	AI223666	2.71	0.027	0.01	
755865	AA496479	2.67	0.073	0.027	
1985448	AI254906	2.63	0.015	0.006	
2292434	AI648563	2.63	0.067	0.026	
810498	AA457154	2.6	0.022	0.009	
2003068	AI223766	2.56	0.041	0.016	
2006294	AI279388	2.54	0.055	0.022	
1985556	AI251321	2.53	0.053	0.021	
1985494	AI254938	2.53	0.058	0.023	
1985040	AI251623	2.52	0.026	0.01	
1984903	AI254586	2.5	0.02	0.008	
2292460	AI648530	2.5	0.07	0.028	
1984941	AI252841	2.5	0.034	0.014	
770747	AA454520	2.48	0.045	0.018	
1985061	AI251642	2.48	0.023	0.009	
2260925	AI591032	2.48	0.085	0.034	
2006359	AI264852	2.47	0.059	0.024	
2292490	AI648545	2.43	0.082	0.034	
770522	AA456762	2.42	0.052	0.022	

CloneID IMAGE#	Accession #	Ratio	Mean Signal 1	Mean Signal 2	Candidate
2292472	AI648533	2.4	0.074	0.031	O1031C
		2.35	0.092	0.039	
2002584	AI223506	2.34	0.024	0.01	
2006508	AI265914	2.34	0.038	0.016	
2002702	AI223706	2.34	0.041	0.018	
2002948	AI223609	2.31	0.028	0.012	
2002540	AI223470	2.31	0.027	0.012	
2260893	AI591004	2.3	0.082	0.036	
2003136	AI250523	2.27	0.023	0.01	
2260680	AI590757	2.26	0.056	0.025	
2219675	AI738841	2.23	0.067	0.03	
2006285	AI279392	2.22	0.033	0.015	
1984994	AI252969	2.18	0.041	0.019	
2260809	AI608710	2.17	0.056	0.026	
2260492	AI590577	2.15	0.081	0.038	
1964644	AI287912	2.14	0.098	0.046	
770961	AA428171	2.13	0.032	0.015	
810067	AA455305	2.12	0.099	0.047	
2107496	AI380369	2.12	0.06	0.028	
2292567	AI648462	2.12	0.085	0.04	
2260591	AI590610	2.11	0.088	0.042	
2261022	AI591055	2.11	0.077	0.036	
2292702	AI612846	2.11	0.088	0.042	
2002858	AI224270	2.09	0.04	0.019	
2006307	AI279408	2.08	0.058	0.028	
1984931	AI252836	2.08	0.035	0.017	
2260917	AI591022	2.08	0.083	0.04	
770927	AA433865	2.07	0.024	0.011	

CloneID IMAGE#	Accession #	Ratio	Mean Signal 1	Mean Signal 2	Candidate
2006292	AI279387	2.07	0.036	0.017	
2261010	AI609587	2.07	0.084	0.041	
741039	AA402296	2.06	0.088	0.043	
2006364	AI264845	2.06	0.013	0.006	
1985021	AI251613	2.04	0.024	0.012	
1374360	AA828959	2.04	0.034	0.017	
1985005	AI252985	2.04	0.046	0.022	
2258189	AI623135	2.03	0.097	0.047	
2006320	AI279404	2.03	0.017	0.008	
770715	AA454497	2.02	0.028	0.014	
2261186	AI609337	2.02	0.064	0.032	
2216400	AI654744	2.02	0.09	0.045	
2006496	AI265908	2.02	0.049	0.024	
2292705	AI612857	2.02	0.089	0.044	
1964750	AI283331	2.01	0.091	0.045	
2260699	AI590770	2.01	0.062	0.031	
1984950	AI252845	2.01	0.045	0.023	
770978	AA428247	2	0.096	0.048	
770998	AA428253	2	0.097	0.048	
2292390	AI648504	2	0.083	0.041	
1985408	AI254597	2	0.045	0.022	
2251282	AI658484	2	0.099	0.049	
Threshold >2.0; Mean Signal 2=0.1-0.2; Rankedby Ratio					
1373633	AA837669	4.24	0.496	0.117	
2122708	AI521444	2.96	0.397	0.134	
753401	AA406401	2.36	0.266	0.113	
2216215	AI654529	2.35	0.336	0.143	
810889	AA459288	2.25	0.316	0.14	

480

CloneID IMAGE#	Accession #	Ratio	Mean Signal 1	Mean Signal 2	Candidate
811573	AA458520	2.19	0.224	0.102	
2216412	AI654749	2.19	0.343	0.157	
754929	AA422036	2.18	0.38	0.174	
981382	AA525940	2.16	0.225	0.104	
1965573	AI364874	2.1	0.305	0.146	
981208	AA526408	2.09	0.322	0.154	
757030	AA429021	2.08	0.344	0.165	
742056	AA405746	2.07	0.314	0.152	
1861875	AI053808	2.07	0.254	0.122	
1373857	AA828779	2.05	0.257	0.125	
1373888	AA828792	2.04	0.223	0.109	
810216	AA463996	2.02	0.377	0.186	
1374124	AA828604	2.02	0.204	0.101	
981025	AA526011	2.01	0.349	0.173	
1212476	AA643476	2	0.217	0.109	
OVM-28		2.63	0.313	0.119	OVM-28
Threshold >2.0; Mean Signal 2>0.2; Ranked by ratio.					
800175	AA581095	2.93	1.213	0.414	
741705	AA402081	2.8	3.921	1.4	
740320	AA477801	2.56	2.479	0.97	
756537	AA436442	2.42	2.437	1.007	
897358	AA496757	2.32	1.246	0.537	
1965502	AI364842	2.26	0.662	0.293	
755590	AA419204	2.25	1.832	0.813	
1964045	AI280624	2.22	0.618	0.279	
1965524	AI364851	2.2	1.354	0.615	
740416	AA477828	2.11	0.477	0.226	
770267	AA434256	2.04	0.48	0.236	

CloneID IMAGE#	Accession #	Ratio	Mean Signal 1	Mean Signal 2	Candidate
811005	AA485364	2.04	0.539	0.264	
741183	AA402548	2.03	2.903	1.431	
727859	AA393566	2.02	0.51	0.252	
OVM-61		2.44	0.818	0.335	OVM-61

Of those sequences presented in Table 7, several that demonstrated good expression profiles were selected for Real Time PCR analysis and their expression examined in a selection of normal and tumor tissues. Table 8 describes the expression profiles for two of these clones.

**Table 8: Real Time PCR Expression Profiles of Sequences Identified
on Ovary Chip #4**

Clone ID	Image #	SEQ ID NO:	Ovary Tumor Tissue and tumor associated tissue	Normal Tissue
1069C	137412 3	7388; 10,894	OE in 6/13	OE in skeletal muscle and brain
O1070C	198544 2	9581; 10,895	OE in 7/13; E in 4/13	E in colon, brain and cerebellum

OE= over expressed; E= expressed

10

All clones subjected to Real Time PCR, in addition to several other clones identified using microarray, were re-sequenced and the sequences were searched against Genbank to identify extended and/or full length sequences. These sequences are disclosed in SEQ ID NO:10,863-10,912.

EXAMPLE 3

PEPTIDE PRIMING OF T-HELPER LINES

Generation of CD4⁺ T helper lines and identification of peptide epitopes derived from tumor-specific antigens that are capable of being recognized by CD4⁺ T cells in the context of HLA class II molecules, is carried out as follows:

Fifteen-mer peptides overlapping by 10 amino acids, derived from a tumor-specific antigen, are generated using standard procedures. Dendritic cells (DC) are derived from PBMC of a normal donor using GM-CSF and IL-4 by standard protocols. CD4⁺ T cells are generated from the same donor as the DC using MACS beads (Miltenyi Biotec, Auburn, CA) and negative selection. DC are pulsed overnight with pools of the 15-mer peptides, with each peptide at a final concentration of 0.25 µg/ml. Pulsed DC are washed and plated at 1×10^4 cells/well of 96-well V-bottom plates and purified CD4⁺ T cells are added at 1×10^5 /well. Cultures are supplemented with 60 ng/ml IL-6 and 10 ng/ml IL-12 and incubated at 37°C. Cultures are restimulated as above on a weekly basis using DC generated and pulsed as above as antigen presenting cells, supplemented with 5 ng/ml IL-7 and 10 U/ml IL-2. Following 4 *in vitro* stimulation cycles, resulting CD4⁺ T cell lines (each line corresponding to one well) are tested for specific proliferation and cytokine production in response to the stimulating pools of peptide with an irrelevant pool of peptides used as a control.

20

EXAMPLE 4

GENERATION OF TUMOR-SPECIFIC CTL LINES USING IN VITRO WHOLE-

GENE PRIMING

Using *in vitro* whole-gene priming with tumor antigen-vaccinia infected DC (see, for example, Yee et al, *The Journal of Immunology*, 157(9):4079-86, 1996), human CTL lines are derived that specifically recognize autologous fibroblasts transduced with a specific tumor antigen, as determined by interferon-γ ELISPOT analysis. Specifically, dendritic cells (DC) are differentiated from monocyte cultures derived from PBMC of normal human donors by growing for five days in RPMI medium containing 10% human serum, 50 ng/ml human GM-CSF and 30 ng/ml human IL-4. Following culture, DC are infected overnight with tumor antigen-recombinant

30

vaccinia virus at a multiplicity of infection (M.O.I) of five, and matured overnight by the addition of 3 µg/ml CD40 ligand. Virus is then inactivated by UV irradiation. CD8+ T cells are isolated using a magnetic bead system, and priming cultures are initiated using standard culture techniques. Cultures are restimulated every 7-10 days using autologous primary fibroblasts retrovirally transduced with previously identified tumor antigens. Following four stimulation cycles, CD8+ T cell lines are identified that specifically produce interferon-γ when stimulated with tumor antigen-transduced autologous fibroblasts. Using a panel of HLA-mismatched B-LCL lines transduced with a vector expressing a tumor antigen, and measuring interferon-γ production by the CTL lines in an ELISPOT assay, the HLA restriction of the CTL lines is determined.

EXAMPLE 5

GENERATION AND CHARACTERIZATION OF ANTI-TUMOR ANTIGEN

MONOCLONAL ANTIBODIES

Mouse monoclonal antibodies are raised against *E. coli* derived tumor antigen proteins as follows: Mice are immunized with Complete Freund's Adjuvant (CFA) containing 50 µg recombinant tumor protein, followed by a subsequent intraperitoneal boost with Incomplete Freund's Adjuvant (IFA) containing 10µg recombinant protein. Three days prior to removal of the spleens, the mice are immunized intravenously with approximately 50µg of soluble recombinant protein. The spleen of a mouse with a positive titer to the tumor antigen is removed, and a single-cell suspension made and used for fusion to SP2/O myeloma cells to generate B cell hybridomas. The supernatants from the hybrid clones are tested by ELISA for specificity to recombinant tumor protein, and epitope mapped using peptides that spanned the entire tumor protein sequence. The mAbs are also tested by flow cytometry for their ability to detect tumor protein on the surface of cells stably transfected with the cDNA encoding the tumor protein.

EXAMPLE 6SYNTHESIS OF POLYPEPTIDES

Polypeptides are synthesized on a Perkin Elmer/Applied Biosystems Division 430A peptide synthesizer using Fmoc chemistry with HPTU (O-
5 Benzotriazole-N,N,N',N'-tetramethyluronium hexafluorophosphate) activation. A Gly-Cys-Gly sequence is attached to the amino terminus of the peptide to provide a method of conjugation, binding to an immobilized surface, or labeling of the peptide. Cleavage of the peptides from the solid support is carried out using the following cleavage mixture: trifluoroacetic acid:ethanedithiol:thioanisole:water:phenol (40:1:2:2:3). After
10 cleaving for 2 hours, the peptides are precipitated in cold methyl-t-butyl-ether. The peptide pellets are then dissolved in water containing 0.1% trifluoroacetic acid (TFA) and lyophilized prior to purification by C18 reverse phase HPLC. A gradient of 0%-60% acetonitrile (containing 0.1% TFA) in water (containing 0.1% TFA) is used to elute the peptides. Following lyophilization of the pure fractions, the peptides are
15 characterized using electrospray or other types of mass spectrometry and by amino acid analysis.

From the foregoing it will be appreciated that, although specific embodiments of the invention have been described herein for purposes of illustration,
20 various modifications may be made without deviating from the spirit and scope of the invention. Accordingly, the invention is not limited except as by the appended claims.

CLAIMS

What is claimed:

- 5 1. A method for detecting the presence of ovarian cancer in a patient, comprising the steps of:
- (a) obtaining a biological sample from a patient;
 - (b) contacting the biological sample with an oligonucleotide that hybridizes to a sequence set forth in any one of SEQ ID NO: 1-10,912 under highly
10 stringent conditions;
 - (c) detecting in the sample an amount of a polynucleotide that hybridizes to the oligonucleotide; and
 - (d) comparing the amount of polynucleotide that hybridizes to the oligonucleotide to a predetermined cut-off value, and therefrom detecting the presence
15 of ovarian cancer in the patient.
2. The method of claim 1, wherein said sequence is selected from a sequence set forth in any one of SEQ ID NO: 10863-10912.
- 20 3. The method of claim 1, wherein said sequence is selected from a sequence set forth in any one of SEQ ID NO: 10864-10869, 10872-10878, 10880-10884 and 10894-10895.
4. The method of claim 1, wherein said detecting in the sample an
25 amount of a polynucleotide that hybridizes to the oligonucleotide is performed by a polymerase chain reaction.
5. The method of claim 1, wherein the biological sample is selected from the group consisting of serum and ovarian tissue.

6. An oligonucleotide useful in the detection of ovarian cancer in a patient, wherein said oligonucleotide hybridizes to a sequence set forth in any one of SEQ ID NO: 1-10,912 under highly stringent conditions.

5 7. A diagnostic kit comprising at least one oligonucleotide according to claim 6.

8. A method for detecting the presence of a cancer in a patient, comprising the steps of:

 (a) obtaining a biological sample from a patient;
10 (b) contacting the biological sample with a binding agent that binds to a polypeptide selected from the group consisting of:

 (i) a polypeptide encoded by a polynucleotide sequence set forth in any one of SEQ ID NO: 1-10,912;

 (ii) a sequence having at least 90% identity to said polypeptide;
15

 (iii) a sequence having at least 95% identity to said polypeptide;

 (c) detecting in the sample an amount of polypeptide that binds to the binding agent; and

20 (d) comparing the amount of polypeptide to a predetermined cut-off value and therefrom detecting the presence of a cancer in the patient.

9. A method for stimulating and/or expanding T cells specific for an ovarian tumor protein, comprising contacting T cells with at least one component
25 selected from the group consisting of:

 (a) a polypeptide sequence selected from the group consisting of:

 (i) a polypeptide encoded by a polynucleotide sequence set forth in any one of SEQ ID NO: 1-10,912;

 (ii) a sequence having at least 90% identity to said polypeptide;
30

- (iii) a sequence having at least 95% identity to said polypeptide;
- (b) a polynucleotide selected from the group consisting of:
- (i) a sequence set forth in any one of SEQ ID NO: 1-10,912;
- 5 (ii) a complement of a sequence set forth in any one of SEQ ID NO: 1-10,912;
- (iii) a sequence consisting of at least 20 contiguous residues of a sequence set forth in any one of SEQ ID NO: 1-10,912;
- (iv) a sequence that hybridizes to a sequence set forth in any
- 10 one of SEQ ID NO: 1-10,912, under highly stringent conditions;
- (v) a sequence having at least 90% identity to a sequence set forth in any one of SEQ ID NO: 1-10,912; and
- (vi) a sequence having at least 95% identity to a sequence set forth in any one of SEQ ID NO: 1-10,912.
- 15
10. An isolated T cell population, comprising T cells prepared according to the method of claim 9.
11. A composition comprising a first component selected from the
- 20 group consisting of physiologically acceptable carriers and immunostimulants, and a second component selected from the group consisting of:
- (a) a polypeptide sequence selected from the group consisting of:
- (i) a polypeptide encoded by a polynucleotide sequence set forth in any one of SEQ ID NO: 1-10,912;
- 25 (ii) a sequence having at least 90% identity to said polypeptide;
- (iii) a sequence having at least 95% identity to said polypeptide;
- (b) a polynucleotide sequence selected from the group consisting of:
- 30 (i) a sequence set forth in any one of SEQ ID NO: 1-10,912;

- (ii) a complement of a sequence set forth in any one of SEQ ID NO: 1-10,912;
- (iii) a sequence consisting of at least 20 contiguous residues of a sequence set forth in any one of SEQ ID NO: 1-10,912;
- 5 (iv) a sequence that hybridizes to a sequence set forth in any one of SEQ ID NO: 1-10,912 under highly stringent conditions;
- (v) a sequence having at least 95% identity to a sequence set forth in any one of SEQ ID NO: 1-10,912;
- (vi) a degenerate variant of a sequence set forth in any one of
10 SEQ ID NO: 1-10,912;
- (c) a T cell population according to claim 10; and
- (d) antigen presenting cells that express a polypeptide selected from the group consisting of:
 - (i) a polypeptide encoded by a polynucleotide sequence set
15 forth in any one of SEQ ID NO: 1-10,912.
 - (ii) a sequence having at least 90% identity to said polypeptide; and
 - (iii) a sequence having at least 95% identity to said polypeptide.

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
6 December 2001 (06.12.2001)

PCT

(10) International Publication Number
WO 01/092581 A3

(51) International Patent Classification⁷: C12Q 1/68,
G01N 33/574, A61K 39/00, C12N 5/06, 5/08

(21) International Application Number: PCT/US01/17756

(22) International Filing Date: 29 May 2001 (29.05.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/207,484 26 May 2000 (26.05.2000) US

(71) Applicant (for all designated States except US): CORIXA
CORPORATION [US/US]; Suite 200, 1124 Columbia
Street, Seattle, WA 98104 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): ALGATE, Paul, A.
[GB/US]; 580 Kalmia Place N.W., Issaquah, WA 98027
(US). HARLOCKER, Susan, L. [US/US]; 7522 13th
Avenue W., Seattle, WA 98117 (US). JONES, Robert
[GB/US]; 900 20th Avenue E., Seattle, WA 98112 (US).

(74) Agents: POTTER, Jane, E., R.; Seed Intellectual Prop-
erty Law Group PLLC, Suite 6300, 701 Fifth Avenue, Seat-
tle, WA 98104-7092 et al. (US).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,
CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH,
GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC,
LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW,
MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK,
SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA,
ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM,
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,
IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF,
CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— with international search report

(88) Date of publication of the international search report:
21 August 2003

(15) Information about Correction:

Previous Correction:

see PCT Gazette No. 15/2002 of 11 April 2002, Section II

For two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.

WO 01/092581 A3

(54) Title: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS OF OVARIAN CANCER

(57) Abstract: Compositions and methods for the therapy and diagnosis of cancer, particularly ovarian cancer, are disclosed. Illus-
trative compositions comprise one or more ovarian tumor polypeptides, immunogenic portions thereof, polynucleotides that encode
such polypeptides, antigen presenting cell that expresses such polypeptides, and T cells that are specific for cells expressing such
polypeptides. The disclosed compositions are useful, for example, in the diagnosis, prevention and/or treatment of diseases, partic-
ularly ovarian cancer.

INTERNATIONAL SEARCH REPORT

 International Application No
 PCT/US 01/17756

A. CLASSIFICATION OF SUBJECT MATTER IPC 7 C12Q1/68 G01N33/574 A61K39/00 C12N5/06 C12N5/08		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) IPC 7 C12Q		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practical, search terms used) EMBL, EPO-Internal		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	DE 198 17 557 A (METAGEN GES FUER GENOMFORSCHUN) 21 October 1999 (1999-10-21) page 1, line 7 - line 39; claims 1-40; figures 1-4; examples 1,2	1,4-11
Y	DATABASE EMBL 'Online! 9 January 1995 (1995-01-09) ADAMS, M.D.: "EST72484 Human Ovary Homo sapiens cDNA" retrieved from EMBL Database accession no. T29202 XP002225474 Cited against SEQ ID NO:1 abstract	1,4-11
-/--		
<input checked="" type="checkbox"/> Further documents are listed in the continuation of box C. <input checked="" type="checkbox"/> Patent family members are listed in annex.		
* Special categories of cited documents : "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier document but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art. "8" document member of the same patent family		
Date of the actual completion of the international search 9 May 2003		Date of mailing of the international search report 22.05.2003
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016		Authorized officer van Klompenburg, W

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 01/17756

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 93 23429 A (LUDWIG INST CANCER RES) 25 November 1993 (1993-11-25)	11
A	page 17, line 26 -page 21, line 7; claim 39; figure 2B	1,4-10
A	EP 0 972 830 A (TAKARA SHUZO CO) 19 January 2000 (2000-01-19) page 9, paragraph 61 -page 10, paragraph 63	1,4-11
X	DATABASE EMBL 'Online! EBI; 13 November 1998 (1998-11-13) STRAUSBERG, ET AL.: "gv54g10.x1 NCI_CGAP_Ov32 Homo sapiens cDNA clone IMAGE:1985442" retrieved from EMBL Database accession no. AI254625 XP002240563 cited in the application Cited against SEQ ID NO: 10895 abstract	1-11

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 01/17756

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
see FURTHER INFORMATION sheet PCT/ISA/210
2. ☐ Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☒ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
1-11, all partially
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☒ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

Invention 1, claims: 1,4-11 all partially

Invention 1 is characterized by the sequence of SEQ ID NO:1.
A method for detecting the presence of ovarian cancer. An oligonucleotide, a diagnostic kit, A method for stimulating/expanding T cells. An isolated T cell population. A composition comprising first and second components.

Invention 2, claims: 1,4-11 all partially

As invention 1 but characterized by the sequence of SEQ ID NO: 2

Inventions 3-10862, claims: 1,4-11, all partially

As invention 1 but where every invention is characterized by one of the sequences of SEQ ID NOs: 3-10862.

Invention 108663, claims: 1,2,4-11 all partially

As invention 1 but characterized by SEQ ID NO: 108663

Invention 10864 claims: 1-11, all partially

As invention 1, but characterized by SEQ ID NO: 10864

Inventions 10865-10869, 10872-10878,10880-10884,
10894-10895, claims: 1-11, all partially

As invention 1 but where every invention is characterized by one of the sequences of SEQ ID NOs: 10865-10869, 10872-10878,10880-10884,10894-10895.

Invention 10870, claims: 1,2,4-11, all partially

As invention 1, but characterized by SEQ ID NO: 10870.

Inventions 10871,10879,10885-10893,10896-10912,
claims: 1,2,4-11, all partially

As invention 1 but where every invention is characterized by one of the sequences of SEQ ID NOs: 10871, 10879,

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

10885-10893, 10896-10912.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.1

Claims 1-5,8 are directed to a diagnostic method practised on the human/animal body, for the search the step of obtaining a sample from a patient has been ignored.

Although claim 9 is directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 01/17756

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
DE 19817557	A	21-10-1999	DE 19817557 A1	21-10-1999
			WO 9953040 A2	21-10-1999
			EP 1291422 A2	12-03-2003
			EP 1073727 A2	07-02-2001
			JP 2002511252 T	16-04-2002
WO 9323429	A	25-11-1993	AU 664752 B2	30-11-1995
			AU 4050593 A	13-12-1993
			WO 9323429 A1	25-11-1993
			CA 2135471 A1	25-11-1993
			EP 0646129 A1	05-04-1995
			JP 7506969 T	03-08-1995
			US 6107472 A	22-08-2000
EP 0972830	A	19-01-2000	AU 6229098 A	09-09-1998
			EP 0972830 A1	19-01-2000
			WO 9837187 A1	27-08-1998

CORRECTED VERSION

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
6 December 2001 (06.12.2001)

PCT

(10) International Publication Number
WO 01/92581 A2

(51) International Patent Classification⁷: C12Q 1/68,
G01N 33/574, A61K 39/00, C12N 5/06, 5/08

(21) International Application Number: PCT/US01/17756

(22) International Filing Date: 29 May 2001 (29.05.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/207,484 26 May 2000 (26.05.2000) US

(71) Applicant (for all designated States except US): CORIXA
CORPORATION [US/US]; Suite 200, 1124 Columbia
Street, Seattle, WA 98104 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): ALGATE, Paul, A.
[GB/US]; 580 Kalmia Place N.W., Issaquah, WA 98027
(US). HARLOCKER, Susan, L. [US/US]; 7522 13th
Avenue W., Seattle, WA 98117 (US). JONES, Robert
[GB/US]; 900 20th Avenue E., Seattle, WA 98112 (US).

(74) Agents: POTTER, Jane, E., R.; Seed Intellectual Prop-
erty Law Group PLLC, Suite 6300, 701 Fifth Avenue, Seat-
tle, WA 98104-7092 et al. (US).

(81) Designated States (national): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,

CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH,
GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC,
LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW,
MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK,
SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA,
ZW.

(84) Designated States (regional): ARIPO patent (GH, GM,
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,
IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF,
CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

- without international search report and to be republished
upon receipt of that report
- with sequence listing part of description published sepa-
rately in electronic form and available upon request from
the International Bureau

(48) Date of publication of this corrected version:
11 April 2002

(15) Information about Correction:
see PCT Gazette No. 15/2002 of 11 April 2002, Section II

For two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.

WO 01/92581 A2

(54) Title: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS OF OVARIAN CANCER

(57) Abstract: Compositions and methods for the therapy and diagnosis of cancer, particularly ovarian cancer, are disclosed. Illustrative compositions comprise one or more ovarian tumor polypeptides, immunogenic portions thereof, polynucleotides that encode such polypeptides, antigen presenting cell that expresses such polypeptides, and T cells that are specific for cells expressing such polypeptides. The disclosed compositions are useful, for example, in the diagnosis, prevention and/or treatment of diseases, particularly ovarian cancer.